

Raimo Lantelankallio

**PELASTAVATKO SUURAINEISTOT
MAAILMAN ILMASTONMUUTOKSELTA?
SUURAINEISTOJEN KÄYTTÄMINEN
ILMASTONMUUTOKSEN TUTKIMISESSA -
KIRJALLISUUSKATSAUS**

TIIVISTELMÄ

LANTELANKALLIO, RAIMO: Pelastavatko suuraineistot maailman ilmastomuutokselta? Suuraineistojen käyttäminen ilmastomuutoksen tutkimisessa - kirjallisuuskatsaus

Diplomityö, 70 sivua

Tampereen Yliopisto

Lokakuu 2019

Tietojohtaminen, Tuote- ja prosessitiedon hallinta

Diplomityö on kirjallisuuskatsaus, jossa käydään läpi 34 kpl alkuvuonna 2019 (tammikuu - kesäkuu) julkaistuja tutkimuksia liittyen avoimesti saatavilla olevien ns. suuraineistojen (big data) käyttämiseen ilmastomuutoksen havainnointiin liittyen. Referoitavat tutkimukset ovat julkaistu aikavälillä tammikuu – kesäkuu 2019.

Ilmastomuutos on hyvin monimutkainen ilmiö, jonka ymmärtäminen on hyvin hankalaa. Yksi näkökulma ilmiön havainnoimiseen on erilaisten suuraineistojen käyttäminen ja näiden aineistojen analysointi. Tarkastellut tutkimukset on haettu Google Scholar -hakupalvelusta.

Ilmastomuutokseen liittyvässä tutkimuksessa erilaisten suuraineistojen käyttäminen on yhä yleisempää. Tyypillisessä suuraineistoa käyttävässä tutkimuksessa on käytetty matemaattisia analyysijä, joiden avulla on tyypillisesti tehty erilaisia mallinnuksia ja simulaatioita. Tyypillisesti ilmastomuutokseen liittyvät ja suuraineistoa käyttävät tutkimukset liittyvät kaupunkisuunnitteluun, erilaisten riskien ennakkointiin, biodiversiteetteihin ja näihin liittyviin muutokseen sekä muihin vastaaviin. Osa tutkimuksista käsittelee pelkästään suuraineistojen hyödyntämistä ilmastomuutokseen liittyvässä tutkimuksessa. Erilaista dataa on käytettävissä nykyään yhä suurempia määriä ja erityisesti datamassoja yhdistämällä voidaan saada uutta tietoa ilmiöstä. Suuraineistojen analyysit eivät todennäköisesti ole yksistään riittäviä vaan analyysien rinnalle tarvitaan erilaisia perinteisempiä tutkimusmetodeja kuten esim. haastatteluja. Lopputyössä referoituja tutkimusten aihepiirejä oli käsitelty myös Ylen Uutisissa, painottuen vuosiin 2018-2019.

Referoiduissa tutkimuksissa käsiteltiin mm. seuraavia teemoja äärisäiden lisääntyminen ja näihin varautuminen, napa-alueiden jääpeitteiden sulaminen ja tähän liittyvät uhkakuvat, metsäpalot, vaikutukset maatalouteen ja biodiversiteetteihin sekä eliöiden sukupuuttoon, kaupunkien varautuminen, ilmastopakolaiset, energiantuotannon ja energiahuollon varautuminen ilmastomuutokseen, maanjäristysten ennakkointi, suuraineistojen määrän kasvu ja hyödyntäminen ja globaalien ilmastomallien ja ennusteiden soveltaminen paikallisesti. Myös suuraineistojen käyttäminen muiden tutkimustapojen (esim. haastattelut) rinnalla näyttäisi olevan mahdollinen tulevaisuuden kehityssuunta, jolloin suuraineistot ovat enempi yksi osa ilmastomuutoksen tutkimuksessa.

Avainsanat: Ilmastomuutos, suuraineistot, big data, kirjallisuuskatsaus, tietojohtaminen

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

ABSTRACT

LANTELANKALLIO, RAIMO: Will Big Data save the World from climate change
– Using Big Data in climate change research – literature review
Master of Science Thesis, 70 pages
University of Tampere
October 2019
Information management, product and process data management

The Master's Thesis is a literature review that reviews 34 studies published in early 2019 (January - June), which use freely accessible big data for climate change related research.

Climate change is a very complex phenomenon whose understanding is very difficult. One aspect of observing the phenomenon is the use of various major data and analysis of these data. The researches examined have been searched from the Google Scholar search service.

In the research on climate change, the use of various major data is increasingly common. Typical big data -related research has used mathematical analyzes that have typically been modeled and simulated. Typically, climate-related big data researches are related to urban planning, prediction of various risks, biodiversity related changes.. Some of the studies deal with big data -related issues relating to climate change. Various big data resources are now available in ever larger amounts, and especially by combining big data sources, new information can be obtained about the phenomenon. Analyzes of big data are unlikely to be sufficient, but more traditional research methods such as interviews are needed alongside analysis. The topics covered in the final thesis were also discussed in Yle News, mainly during 2018-2019.

Referenced studies included: The following themes were noticed from referenced scientific papers: the increase of extreme weather, the melting of polar ice covers and related threats, forest fires, impacts on agriculture and biodiversity, extinction of organisms, preparedness for cities, climate refugees, preparedness of energy production and energy supply, prediction of earthquakes, growth and usage of big data. The usage of global climate models and predictions locally. Also, the use of large data alongside other research methods (eg interviews) seems to be a potential future trend. Big data related research is therefore only one part of climate change -related research.

Keywords: climate change, big data, literature review, information management

The originality of this thesis has been checked using the Turnitin OriginalityCheck service.

ALKUSANAT

Diplomityön tekeminen oli ihan mielenkiintoinen prosessi. Kirjoittelin DI-työtä pikkuhiljaa muiden aktiviteettien ohessa vuonna 2019. Vaikka yrityksistä ei löytynyt aiheita, oli omasta aiheesta kirjoittaminen ihan mielenkiintoista ja sujuvaa. Olen kiinnostunut ilmastonmuutoksesta ilmiönä ja olen opiskellut tietojohdamista ja data-analytiikkaa. Näistä syistä aihevalinta oli melko helppo.

Kiinnostukseni ilmastonmuutosta kohtaan alkoi luettuani Risto Isomäen fiktiivisen "Sarasvatin Hiekkaa" -kirjan ehkä joskus ennen vuotta 2010. Diplomityötä on kirjoitettu kesällä 2019, jolloin Suomessa oli jälleen ennätyshelteet ja diplomityön aihekin alkoi tuntua entistä ajankohtaisemmalta. Kieltämättä diplomityötä tehdessä tulin itsekkin välillä miettineeksi, että millaisessa maailmassa omat lapseni mahtavat varttua. Ehkä itse ajattelen asiasta niin, että meidän kannattaisi nyt valmistautua erilaisiin muutoksiin, kuten diplomityössä pikaisesti referoidussa John Michael Greerin "The Long Descent" -kirjassakin pohditaan. Ehkä ilmastonmuutokseen pitäisi ensisijaisesti reagoida asennemuutoksella ja luopua lyhytkestoiseen ylikuluttamiseen perustuvasta kulttuuristamme. Tyypillisesti kuitenkin ilmastonmuutoksen hillitsemiseksi yritetään kuitenkin kehittää erilaisia teknologisia ratkaisuja, joilla ehkä enemmänkin pyritään jatkamaan nykyistä menoa mahdollisimman pitkään. Erilaiset meitä edeltävät yhteiskunnat olivat huomattavasti omaamme yksinkertaisempia ja käytetty teknologia oli omaamme yksinkertaisempaa. Kuitenkin monet näistä yhteiskunnista lopulta romahtivat, tyypillisesti satojen vuosien aikana. Toivokaamme ettei oma sivilisaatiomme kulje samoja kehityspolkuja. Ja mikäli kulkee, ehkä juuri nyt on hyvä valmistua eri tavoin.

DI-työstä tuli jonkinlainen poikkitieteellinen *smörgåsbord*, mutta toisaalta ilmastonmuutos on myös hyvin monimutkainen haaste, joka vaatii nimenomaan poikkitieteellistä ajattelua.

Haluan kiittää DI-työn ohjaajaa, professori Hannu Kärkkäistä ja professori Samuli Pekkola tuesta ja ymmärryksestä. Lisäksi haluan kiittää perhettäni tuesta DI-työn tekemisessä. Kiitokset kuuluvat myös hollantilaiselle vaihtoehtomusiikkiyhtye "Omnia"lle, jonka musiikin avulla pysyin itsekkin iskussa DI-työtä kirjoittaessa.

*"The smell of sun on autumn leaves,
He said take good care of all you see,
Praise the earth down on your knees"
(Omnia, "Old Man Tree").*

Tampereella 3.10.2019

Raimo Lantelankallio

SISÄLLYS

1	JOHDANTO.....	1
1.1	Tutkimuksen tausta ja lähtökohdat.....	1
2	METODOLOGIA JA METODIT.....	3
2.1	Tutkimuskysymysten asettaminen.....	3
2.2	Tietokantojen, hakutermien ja seulan valinta.....	4
2.3	Metodologisen seulan asettaminen ja katsauksen suorittaminen.....	5
2.4	Synteesin tekeminen.....	7
3	ILMASTONMUUTOS JA SUURAINEISTOT YLEISELLÄ TASOLLA.....	9
3.1	Ilmastonmuutos tutkimuskohteena.....	9
3.2	Suuraineistojen 4V-malli.....	9
3.3	Ilmastonmuutokseen liittyvät suuraineistot.....	10
3.4	Mittaustiedot suuraineistoina.....	11
3.5	Sosiaalinen media suuraineiston lähteenä.....	14
3.6	Muita kuin suuraineistojen käyttöön perustuvia ilmastonmuutostutkimuksia.....	18
3.7	Aiempia diplomitöitä ja muita julkaisuja aiheesta Suomessa.....	19
3.7.1	Aiempia suuraineistoihin liittyviä tutkimuksia.....	19
3.7.2	Aiempia ilmastonmuutosta ja suuraineistoja yhdisteleviä tutkimuksia.....	20
4	TULOKSET.....	23
4.1	Tutkimuksia liittyen käytettäviin data-aineistoihin.....	23
4.2	Riskienhallintaan liittyviä tutkimuksia.....	34
4.3	Tyypillisiä aluespesifisiä tutkimuksia.....	39
5	JOHTOPÄÄTÖKSET.....	47
5.1	Referoidut tutkimukset ja 4V-malli.....	47
5.2	Suuraineistojen hyödyntämisen malli.....	49
5.3	Johtopäätösten yhteenvetoa ja analysointia.....	51
5.4	Tutkimuskysymysten vastaukset.....	60
5.4.1	Kirjallisuustutkimuksen merkitys ja löydöt.....	66
5.5	Suuraineistot ilmastonmuutoksen jälkeen?.....	67
5.5.1	Suuraineistojen sukupuuttoaalto -skenaario.....	67
5.5.2	Suuraineistojen integroituminen -skenaario.....	69
5.5.3	Poikkitieteelliset suuraineistot -skenaario.....	69
	LÄHTEET.....	71

LYHENTEET JA MERKINNÄT

Termi	Selitys
Avoim data	Vapaasti kenen tahansa saatavilla oleva suuraineisto.
Big Earth Data	Aineisto, joka koostuu tyypillisesti erilaisista mittaustuloksista liittyen maapallosta mitattuihin asioihin (esim. lämpötilat eri puolilla maapalloa eri aikoina).
Big data	Aineisto, jossa on hyvin paljon erilaista tietoa. Tyypillisesti suuraineisto voi sisältää paljon myös hyödytöntä, vanhentunutta, epäoleellista tai epätarkkaa tietoa. Suuraineiston tietoja analysoimalla, yhdistelemällä ja suodattamalla voidaan tuottaa uutta tietoa.
Crowdsourcing	Joukkoistaminen, ongelmanratkaisutapa jossa hyödynnetään esim. suuren ihmisjoukon päätöksentekoa
Data monetization	Liiketoimintaa, jossa joko myydään dataa tai data-analyysien perusteella tuotettua palvelua tai tietoa.
Downscaling	Termi, joka tarkoittaa globaalin mallin (esim. ilmastomalli) muuntamista siten, että mallia voidaan käyttää paikallisten muutosten mallintamiseen.
IoT	Internet of Things, internetiä käyttävien laitteiden muodostama verkko
IPCC	Intergovernmental Panel on Climate Change. Hallitustenvälinen ilmastomuutospaneeli. Ilmastotieteen asiantuntijoista koottu tutkimusryhmä.
Metadata	Erilaista tietoa itse dataan liittyen
NASA	National Aeronautics and Space Administration, Yhdysvaltojen ilmailu- ja avaruushallintovirasto.
Some	Sosiaalinen media (esim. Twitter, Facebook).
SSAB	Ruotsalainen, raakateräksen tuotantoon erikoistunut yritys
Suuraineisto	Big datan suomennettu termi.
Upscaling	Termi, joka tarkoittaa paikallisen mallin (esim. ilmastomalli tai muu vastaava) soveltamista globaalissa mittakaavassa tai mallin kopioimista muualle.

1 JOHDANTO

Tässä diplomityössä on pyritty selvittämään, miten erilaisia suuraineistoja käytetään ilmastonmuutokseen liittyvässä tutkimuksessa. Tämä diplomityö on kirjallisuustutkimus.

Ilmastonmuutosta pidetään yhtenä merkittävimmistä tulevaisuuden riskeistä ihmiskunnalle. Erilaiset ilmastonmuutoksen aiheuttamat riskit, kuten ilmastonmuutokseen sopeutumisen epäonnistuminen sekä erilaiset äärisäät on määritelty World Economic Forumin raportissa ihmiskunnan suurimmiksi uhiksi (World Economic Forum 2019). Toinen tyypillinen nykyajan ilmiö on erilaisen digitaalisen tiedon määrän kasvaminen ja erilaisten sensorien yleistyminen. Tietoja on saatavilla yhä enemmän ja yhä helpommin.

Suuraineistolle ei ole yhtä ainutta määritelmää. Suuraineistona voidaan esim. pitää mitä tahansa tietomäärää, mikä on liian suuri ihmisen ymmärrettäväksi sellaisenaan. Vastaavasti ei-suuraineisto voi olla esim. varsinaiselle suuraineistolle tehty visualisointi, analysointi tai muu prosessointi, joka on muuttanut suuraineiston sisältämän tiedon sellaiseksi, että ihminen kykenee tiedon ymmärtämään.

1.1 Tutkimuksen tausta ja lähtökohdat

Tämä diplomityö on tehty pelkästään tekijän oman mielenkiinnon vuoksi. Diplomityössä on käyty läpi alkuvuonna 2019 (tammikuu - kesäkuu) julkaistuja Googlesta ja erityisesti Google Scholar -palvelusta löytyviä tutkimuksia, joissa erilaisia suuraineistoja on käytetty ilmastonmuutoksen tutkimukseen liittyen. Työn tavoitteena onkin ollut lähinnä diplomityön tekijän tiedon lisääminen eikä siis esim. lisäarvon tuottaminen jollekin kaupalliselle yritykselle. Koska aihepiiri on hyvin laaja, on aihepiiriä rajoitettu keskittymällä kuluvana vuonna (2019) julkaistuihin tutkimuksiin.

Tarkasteltavissa tutkimuksissa on käytetty tyypillisesti vapaasti saatavilla olevia suuraineistoja. Tämänäyttelisten (hyvin laajojen ja mahdollisesti ei-kaupallisten) suuraineistojen kerääminen ei ole välttämättä taloudellisesti kannattavaa tai edes mahdollista esim. yksityisille yrityksille. Tyypillinen ilmastonmuutokseen liittyvä suuraineisto voi olla esim. erilaiset lämpötilatiedot, sademäärät ja muut vastaavat tiedot lukuisilta mittausasemilta kymmenien vuosien ajalta. Tämänäyttelistä mittausdataa kerää tyypillisesti valtiolliset toimijat ja näitä suuraineistoja on vapaasti saatavilla. Avoimen datan suuraineistoista pystytään myös tekemään tarvittaessa vertailevia tutkimuksia, mikäli olemassa olevien tutkimusten luotettavuutta halutaan arvioida. Luonnontieteellisen datan lisäksi voidaan käyttää myös erilaisia sosiaalisesta mediasta saatuja tietoja. Sosiaalisen median data ei kuitenkaan välttämättä ole kaikilta osin vapaasti saatavilla. Kuitenkin sosiaalisen median kautta saatavat suuraineistot voivat

tuoda lisää näkemystä ilmastonmuutokseen liittyvään tutkimukseen, on tämän tyyllisiä tutkimuksia käsitelty tässä diplomityössä soveltuvilta osin.

Tyypillisesti tutkimuksissa käsitellään suuria aineistoja ja rajanveto sille, mikä pidetään suuraineistona (big data) ja mitä tavallisena suurena aineistona, on häilyvä. Tarkasteltavat tutkimukset on valittu sillä perusteella, että mikäli tutkimuksen tekijät käyttävät aineistostaan termiä suuraineisto, on kyseisiä tutkimuksia referoitu.

Tutkimuksia käsitellessä on pyritty myös etsimään tutkimuksia sivuavia uutisia keskittyen vuosiin 2018-2019. Uutisia on etsitty pääosin Ylen sivuilta ja uutisten etsinnän tarkoitus on saada parempi käsitys siitä, miten tutkimuksiin liittyviä asioita ja ilmiöitä on käsitelty mediassa.

2 METODOLOGIA JA METODIT

Diplomityön tyyppi on puhdas kirjallisuustutkimus, jossa on käyty läpi Google Scholar -palvelusta saatavilla olevia vuonna 2019 julkaistuja tutkimuksia. Tiedonkeruuta on tehty myös Googlen avustuksella sekä uutisia seuraamalla. Ilmastomuutos on esiintynyt viime vuosina yhä useammin erilaisissa uutisissa ja uutisten avulla voidaan saada lisänäkökulmaa tutkimuksiin.

Diplomityössä ei käsitellä kovin tarkasti itse avointa dataa yleisellä tasolla (määritelmiä tms.) vaan keskitytään avoimen datan käyttämiseen ilmastomuutokseen liittyvässä tutkimuksessa erityisesti vuonna 2019 julkaistuissa tutkimuksissa. Avoimen datan määritelmiä ja hyödyntämistä yleisellä tasolla on käsitelty tämän diplomityön lähteissä (esim. Heinonen 2016 sisältää hyvän koosteen).

Tämä diplomityö ei anna yksiselitteistä vastausta sille, miten suuraineistoja olisi hyödynnettävä ilmastomuutokseen liittyvässä tutkimuksessa, vaan diplomityö on pelkästään yleiskatsaus nykyiseen tilanteeseen. Diplomityön avulla ei voi luoda uutta liiketoimintaa.

Diplomityössä on tarkasteltu myös tutkimuksia, joita voidaan soveltaa vaivatta myös ilmastomuutokseen liittyviä asioita tutkittaessa. Tällaisia tutkimuksia ovat tyypillisesti ilmastoon, ekosysteemeihin ja muihin maapallosta mitattaviin asioihin liittyvät tutkimukset ja suuraineistot.

Kirjallisuustutkimuksessa on noudatettu soveltuvilta osin Arlene Finkin kirjallisuustutkimuksen mallia (Fink 2019 sekä mukailen Okoli 2010).

2.1 Tutkimuskysymysten asettaminen

Diplomityön tutkimuskysymyksenä on:

- Miten erilaisia suuraineistoja on käytetty tässä lopputyössä tarkastelluissa ilmastomuutokseen liittyvässä tutkimuksessa vuonna 2019 julkaistuissa tutkimuksissa?

Alikysymyksiä ja tutkimuksen pääkysymystä täsmentävinä kysymyksiä ovat seuraavat kysymykset:

- Millä tavalla suuraineistot auttavat ilmastomuutokseen liittyvien haasteiden ymmärtämisessä?
- Minkälaisia keinoja suuraineistot tarjoavat ilmastomuutokseen liittyvässä tutkimuksessa verrattuna pienaineistoihin?
- Miksi suuraineistoja käytetään ilmastomuutokseen liittyvässä tutkimuksessa?
- Näkyykö Suomen Ylen vuonna 2019 julkaistuissa uutisissa vastaavia teemoja ja

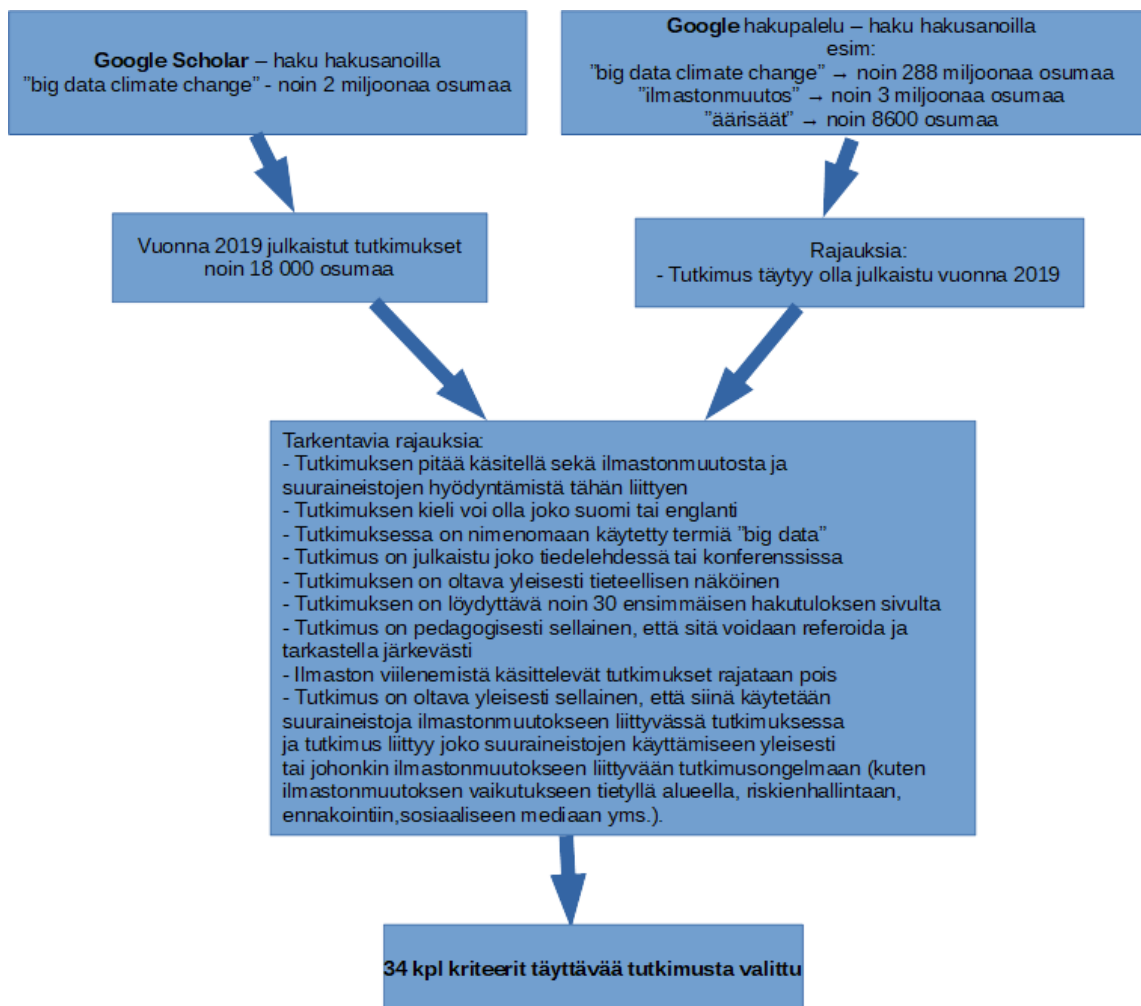
asioita kuin käsitellyissä tutkimuksissa?

- Pelastavatko suuraineistot meidät ilmastonmuutokselta – mikä on suuraineistojen rooli nykyisessä ilmastonmuutokseen liittyvässä tutkimuksessa?

2.2 Tietokantojen, hakutermien ja seulan valinta

Suuraineistoja on vapaasti saatavissa erilaisista sosiaalisen median lähteistä, sanomalehdistä sekä julkisista lähteistä (kuten Tilastokeskuksen julkaisut Suomessa). Lisäksi erilaisilla mittalaitteilla voidaan kerätä suuria määriä dataa, joka saattaa olla ainoastaan datan kerääjien yksityisessä käytössä. Diplomityössä on tarkasteltu erilaisia tutkimuksia, joissa eri tavoin kerättyjä suuraineistoja on hyödynnetty ilmastonmuutoksen tutkimuksessa. Julkisesti saatavilla olevat tutkimukset hyödyntävät tyypillisesti avoimesti saatavilla olevia suuraineistoja koska eri yritysten keräämät aineistot ovat usein liikesalaisuuksia. Tässä diplomityössä ei ole eroteltu julkisia ja yksityisiä suuraineistoja eikä niiden hyödyntämistä.

Hakusanoina Google Scholar -palvelussa on käytetty sanoja "climate change" ja "big data" sekä näiden yhdistelmiä. Lisäksi tutkimuksia on haettu myös käyttäen Googlen normaalia hakukonetta. Hakusanoina on ollut "ilmastonmuutos", "big data", "lopputyö", "tutkimus" sekä aihepiiriin liittyviä hakusanoja kuten "äärisäät".



Kuva 1. Referoitavien tutkimusten valinta (lukumäärät haettu 3.9.2019).

2.3 Metodologisen seulan asettaminen ja katsauksen suorittaminen

Referoitavat tutkimukset on jaettu kolmeen kategoriaan. Jako on karkea ja perustuu lopputyön tekijän omaan näkemyskseen ja osan tutkimuksista voisi todennäköisesti jakaa yhtä hyvin useampaan kuin yhteen kategoriaan. Osa tutkimuksista käsittelee suuraineistojen käyttämistä yleisesti ilmastonmuutokseen liittyvässä tutkimuksessa. Lisäksi on etsitty erityisesti riskinhallintaan liittyviä tutkimuksia sekä tiettyyn tarkasti määriteltuihin maantieteellisiin alueisiin liittyviä tutkimuksia.

Tutkimuksessa tarkastellut tutkimukset on rajattu vuonna 2019 julkaistuihin tutkimuksiin (tammikuu - kesäkuu 2019) koska tarkoitus on ollut ainoastaan saada jonkinlainen käsitys tämänhetkisestä tutkimuksesta ja siitä, minkälaisia tutkimuksia ja mistä aihepiireistä on viime aikoina tehty. Kasvava kiinnostus ilmastonmuutokseen on johtanut siihen, että erilaisia ilmastonmuutoksia ja suuraineistoja käsitteleviä tutkimuksia on tehty lukuisia, ja kaikkien näiden läpikäyminen ja analysointi olisi hyvin työlästä. Lisäksi suuraineistojen käsittely ja tähän liittyvät teknologiat kehittyvät nopeasti, jolloin vanhempien tutkimusten tarkastelu ei ehkä olisi relevanttia.

nykytilanteen selvittämisen näkökulmasta. Lisäksi kun tämän tutkimuksen yhtenä tutkimuskysymyksenä on selvittää referoitujen tutkimusten näkyvyyttä Ylen uutisissa, niin tarkasteltavien tutkimusten määrää on pyritty rajoittamaan selkeyden vuoksi.

Suuraineistojen tarkastelussa on käytetty 4V-mallia (IBM 2019), jota käsitellään tarkemmin tässä dokumentissa kappaleessa 2.2. Jotta suuraineiston käyttämistä voidaan tarkastella, suuraineiston on oltava tarpeeksi soveltuva 4V-malliin. Tällöin suuraineistotutkimuksen tarkastelu on pedagogisessakin mielessä järkevää. On huomattava, että osa referoiduista tutkimuksista käsitteli suuraineistojen käyttöä ilmastomuutokseen liittyvässä tutkimuksessa ilman että käytettävissä olisi ollut jokin tietty suuraineisto, jota olisi analysoitu jonkun tietyn ongelman ratkaisemiseksi. Tämän tyyppiset tutkimukset liittyivät enempi ilmastomuutoksen data-analytiikan haasteita tutkimukseen.

Google Scholar on valittu käytettäväksi siksi, koska palvelu tarjoaa kattavan määrän erilaisia tutkimuksia ja Google Scholarin hakutoiminnot ovat helppoja ja kyseinen hakukone on yleisesti käytetty.

Valittuihin tutkimuksiin perehdytään tutkimuskysymysten kautta ja tutustutaan suuraineistojen käyttämiseen ilmastomuutokseen liittyvässä tutkimuksessa.

Diplomityön tarkoitus on toimia yleiskatsauksena suuraineistojen käyttöön ilmastomuutokseen liittyen. Tästä syystä diplomityössä ei mennä kovin syvälle matemaattisiin ja teoreettisiin data-analyysin menetelmiin vaan keskitytään nimenomaan lähinnä tarkastelemaan vuonna 2019 julkaistuja tutkimuksia aiheeseen liittyen. Vastaavasti tässä diplomityössä ei myöskään käsitellä ajanpuutteen vuoksi data-analytiikan tai suuraineistojen peruskäsitteitä, vaan näiden oletetaan olevan lukijalla jo ennestään tuttuja. Mikäli olisi haluttu tutkia esim. 50 000 tutkimusta, olisi todennäköisesti ollut järkevintä analysoida automaattisesti tutkimusten yhteenvetoja (abstract) ja laskea erilaisia tilastollisia arvoja. Koska lopputyön tarkoitus oli kuitenkin tarkastella suuraineistojen hyödyntämistä esimerkkitapauksissa, tutkimusten määrä oli pakko rajoittaa pieneksi. Vastaavasti ottamalla tarkasteluun tutkimuksia pitkältä aikaväliltä, olisi saatu kattavampi kuva erilaisista trendeistä. Trendejä voivat olla esim. ilmastomuutoksen tutkimuksen painopisteiden (maanviljely, energiantuotanto, myrskyjen ennustaminen tms.) kehityksen hahmottaminen. Toisaalta tämäntyylisiä tutkimuksia on todennäköisesti tehty jo aiemmin ja tällöin tutkittavaksi saattaisi jäädä lopulta vain lyhyen aikavälin (viime vuodet) tarkastelu. Alla olevassa taulukossa on hahmoteltu lyhyen ja pidemmän aikavälin kirjallisuustutkimusten eroja.

Lyhyen aikavälin (2019 alkuvuosi) kirjallisuustutkimus	Pidemmän aikavälin kirjallisuustutkimus
Tarkempi analysointi tutkimusten suuraineistojen käytöstä	Eri käyttökohteiden trendien havainnointi
Tutkimusten ja uutisoinnin tarkempi vertailu sekä analysointi	Numeerinen ja tilastollinen analyysi
Yleiskuva tietyn hetken tutkimuksen aiheista ja painopisteistä	Laajempi kuva suuraineistojen käytöstä

Lyhyen aikavälin (2019 alkuvuosi) kirjallisuustutkimus	Pidemmän aikavälin kirjallisuustutkimus
Nykyisin käytössä olevien työkalujen (esim. Google Earth Engine) selvittäminen	Laajempi kuva työkalujen käyttötrendeistä
Rajallinen määrä tarkasteltavia tutkimuksia	Tarkasteltavia tutkimuksia erittäin suuri määrä

Taulukko 1. Lyhyen ja pitkän aikavälin kirjallisuustutkimusten toteutuksen vertailua.

Google Scholarin hakusanalla "climate change" löytyy noin 118 000 tutkimusta (pois lukien patentit ja lainaukset). Tässä lopputyössä on käsitelty ja tutkittu ainoastaan 34 tutkimusta, ja tämän lopputyön tarkoitus ei ole antaa yleiskuvaa ilmastonmuutoksen tutkimuksesta, vaan tarkastella ainoastaan hyvin pientä osajoukkoa tietyyppisistä tutkimuksista. Ilmastonmuutokseen liittyvästä tutkimuksesta on tehty aiemmin kirjallisuustutkimuksia, jossa on käsitelty sekä suuraineistoja että myös pienaineistoja käytäviä tutkimuksia (kooste tällaisesta tutkimuksesta on esim. Haunschild et al. 2016). Erilaisia suuraineistoja on käytetty myös aiemmin tehdyissä ilmastonmuutokseen liittyvissä tutkimuksissa ja keskittymällä alkuvuoteen 2019, on haluttu saada yleiskuva nimenomaan nykyisestä tilanteesta ja mahdollisista tämän hetken kehityspoluista. Lisäksi kattavampi tekstuaalinen analyysi pidemmän aikavälin tutkimuksista olisi helposti paisunut satojen sivujen mittaiseksi.

Suuraineistoanalyysit ja tietotekniset työkalut kehittyvät nopeasti, jolloin rajoittamalla katsaus viimeisimpiin tutkimuksiin, saadaan myös paremmin näkemystä suuraineistoanalyysien tekemisestä nykyisin työkaluin. Lisäksi diplomityön aiheena on tutustuminen suuraineistojen käyttäminen ilmastonmuutokseen liittyvässä tutkimuksessa eikä niinkään tilastollinen analyysi esim. pitkän ajan tutkimusten trendeistä.

Lisäksi erilaiset ilmastonmuutokseen liittyviä asioita on uutisoitu laajasti vuonna 2019 ja uutisten sekä tutkimusten sisällöllinen vertailu voi antaa uutta näkökulmaa ilmastonmuutoksesta käytävään keskusteluun.

Ilmastonmuutokseen liittyvästä tutkimuksesta on helposti saatavissa erilaista tietoa pelkästään Google Scholarin hakutulosten lukumääriä vertailemalla. Sen sijaan tutkimusten tarkempi analyysi vaatii kuitenkin syvällisempää perehtymistä tutkimuksiin ja muuhun taustamateriaaliin.

2.4 Synteesin tekeminen

Lopuksi, kun suuraineistojen analysointi on saatu valmiiksi, on saadusta tiedosta muodostettu jonkinlainen yleiskuva tutkimuksen nykytilanteesta sekä arvioitu erilaisia mahdollisia tulevaisuuden kehityspolkuja.

Lisäksi analysoiduista tutkimuksista pyritään löytämään yleisellä tasolla jonkinlainen malli sille, miten suuraineistoja kannattaisi käyttää ilmastonmuutokseen liittyvässä tutkimuksessa referoidun aineiston perusteella. Koska tarkasteltavien tutkimusten määrä on pieni (34 kappaletta), malli on lähinnä suuntaa antava.

Laaduntarkkailu on varmistettu lähinnä sillä, että käytetyt tutkimukset ovat konferenssijulkaisuja, tieteellisissä lehdissä julkaistuja tai muuten diplomityön tekijän käsityksen mukaan luotettavia.

Tämä diplomityö ei anna yksiselitteistä vastausta sille, miten suuraineistoja olisi hyödynnettävä ilmastonmuutokseen liittyvässä tutkimuksessa, vaan diplomityö on pelkästään yleiskatsaus nykyiseen tilanteeseen. Diplomityön avulla ei voi luoda uutta liiketoimintaa.

3 ILMASTONMUUTOS JA SUURAINEISTOT YLEISELLÄ TASOLLA

Ilmastonmuutokseen liittyvässä tutkimuksessa käytetään tyypillisesti erilaisista mittaustiedoista koostuvia suuraineistoja. Näiden lisäksi nykyään on mahdollista hyödyntää myös esim. sosiaalisesta mediasta kerättyjä suuraineistoja ilmastonmuutokseen liittyvässä tutkimuksessa. Tässä kappaleessa tarkastellaan suuraineiston määritelmää sekä erityisesti ilmastonmuutokseen liittyviä suuraineistoja.

3.1 Ilmastonmuutos tutkimuskohteena

Ilmastonmuutoksen yksi näkyvimmistä toimijoista on IPCC (Intergovernmental Panel on Climate Change) joka koostuu eri maiden tutkijoista ja asiantuntijoista. Tässä dokumentissa ei käsitellä ilmastonmuutokseen liittyviä syitä mutta netistä löytyy vapaasti saatavilla pohjatietoja ilmiöstä (esim. IPCC 2004).

Vuonna 2016 julkaistussa tutkimuksessa on analysoitu 11 944 tutkimuksen tiivistelmiä (abstract) ja tutkimuksen mukaan 97 prosentissa tutkimuksissa ilmaston lämpenemistä pidetään ihmisen aikaansaamana (Cook et al. 2016).

Ilmastonmuutokseen liittyviä tutkimuksia on julkaistu lukuisia. Vuosina 1980-1990 ilmastonmuutokseen liittyviä tutkimuksia julkaistiin noin alle tuhat vuodessa kunnes ilmiötä alettiin tutkia enemmän. Ajanjakson 1990-2000 lopussa ilmastonmuutokseen liittyviä tutkimuksia julkaistiin jo 5000 vuodessa ja vuonna 2014 vastaavia tutkimuksia on julkaistu yli 2500 vuodessa. Pääosin (tutkimusten lukumäärän mukaan) tutkimukset liittyvät luonnontieteisiin (natural sciences). Muita merkittäviä tieteenaloja ovat insinööritieteet (engineering), sosiaalitieteet (social sciences) sekä maatalous (agriculture) (Haunschild et al. 2016).

3.2 Suuraineistojen 4V-malli

4V-malli on yksi määrittely sille, millaiset aineistot luokitellaan suuraineistoiksi. 4V mallissa suuraineistoilla voi olla yksi tai useampi 4V-mallissa esitetyistä ominaisuuksista, joita ovat (mukaillen IBM 2019):

- Volume, eli tiedon suuri määrä
 - Esim. puhelinoperaattorien keräämät tiedot kaikista matkapuhelinten käytöstä, erilaiset automaattisesti generoituvat ja päivittyvät datamäärät kuten esim. satelliittikuvat ja säätiedot
- Velocity, eli tietomassan nopea muuttuminen
 - Esim. pörssikaupan kaupankäyntitiedot, lentokoneen tai auton sensorien

reaaliajassa keräämät tiedot

- Variety, eli tietomassan sisältämän tiedon vaihtelevuus
 - Esim. terveydenhuollon potilastiedot, sosiaalisen median aineistot
- Veracity, eli tietomassan tietojen totuudenmukaisuus
 - Esim. tietomassa voi sisältää runsaasti tietoa, jonka hyödyllisyys ja hyödynnettävyys voi olla kyseenalainen tai tuntematon.

Lisäksi suuraineistojen sisältämää dataa pitää pystyä hyödyntämään. Suuraineistojen määritelmä on häilyvä ja muuttuu teknologian ja tallennuskapasiteetin muuttuessa. Ilmastonmuutokseen liittyvässä tutkimuksessa tyypillisiä suuraineistoja ovat esim.

- Pitkän ajan kuluessa automaattisesti kerätty valtava satelliittikuva-aineisto, jossa voi olla esim. infrapunakameralla ja tavallisella kamaralla otettuja valokuvia sekä muuta mittaustietoa. Tyypillisesti tällainen tietomäärä voi olla niin suuri, että sen analysointi käsin on mahdotonta (4V-mallin Volume).
- Pitkän aikavälin kuluessa kerätyt säätiedot, eri lajien kasvinäytteet, erilaiset säähavainnot (4V-mallin Variety).
- Sosiaalisesta mediasta ja uutisista automaattisesti kerätyt ja koostetut aineistot (4V-mallin Variety).

Esimerkkien suuraineistojen manuaalinen käsittely on käytännössä mahdotonta eikä välttämättä voida tietää, että onko suuraineistoissa hyödyllistä tietoa ja jos on, missä ja miten tämän tiedon saa esiin (4V-mallin Veracity).

Ilmastonmuutosta tutkittaessa käytössä on tyypillisesti tietoa pitkältä aikaväliltä ja tieto voi koostua eri tavoin hankituista (osin automaattisesti ja osin manuaalisesti) hankituista tiedoista.

3.3 Ilmastonmuutokseen liittyvät suuraineistot

Esimerkkinä ilmastonmuutokseen liittyvästä tutkimuksesta voi olla vaikka jonkun tietyn jäätikön tutkimus. Mahdollisia tutkimustapoja ovat esim.

- Jäätiköllä ja laboratoriossa tehtävät mittaukset
- Jäätikön luona asuvien sekä asiantuntijoiden haastattelut
- Mahdollisten digitaalisten suuraineistojen (esim. satelliittikuvat, säätiedot) analysointi
- Mahdollisten jäätiköltä aiemmin kerättyjen (ei-digitaalisten) näytteiden analysointi
- Aiempien vastaavien tutkimusten analysointi

Tiettyyn jäätikköön liittyvä tutkimus voi olla osana jotain suurempaa tutkimusta, johon saatetaan tarvita lisää tietoa. Tutkimusryhmällä ei välttämättä ole mahdollisuutta itse vierailla jäätiköllä ja tutkimusta saatetaan haluta myös automatisoida mahdollisimman pitkälle. Mikäli tutkittavaan jäätikköön liittyy jotain vapaasti saatavilla olevia (tutkimusryhmällä ei välttämättä ole resursseja kerätä itse suuraineistoa) suuraineistoja, niin näiden hyödyntäminen voi olla järkevää. Tyypillisesti erilaiset suuraineistot on

kerätty esim. vuosikymmenten ajalta, jolloin ilmaston muuttumista voidaan havainnoida pitkältä aikaväliltä. Tyypillisesti suuraineistojen sisältämät tiedot saattavat muuttua nopeasti. Tyypillisesti ilmastonmuutokseen liittyvät suuraineistot eivät välttämättä muutu (pois lukien mahdolliset reaaliaikaiset analyysit), vaan aineistojen sisältämä data saattaa sisältää tietoa pidemmältä ajalta, koska hitaita muutoksia ei välttämättä muuten voi havainnoida.

Tyypillistä suuraineistojen käsittelyssä on myös erilaisten visualisointien käyttäminen analyysien tuloksia esitettäessä. Edellä mainitussa jäätikköesimerkissä visualisointi voisi olla vaikka animaatio siitä, miten jäätikkö on muuttunut historian aikana ja miten mahdollisesti jäätikkö muuttuu tulevaisuudessa. Erilaisia ilmastonmuutokseen liittyviä visualisointeja löytyy mm. Atlas for the End of the World -sivustolta (Atlas for the End of the World 2019).

3.4 Mittaustiedot suuraineistoina

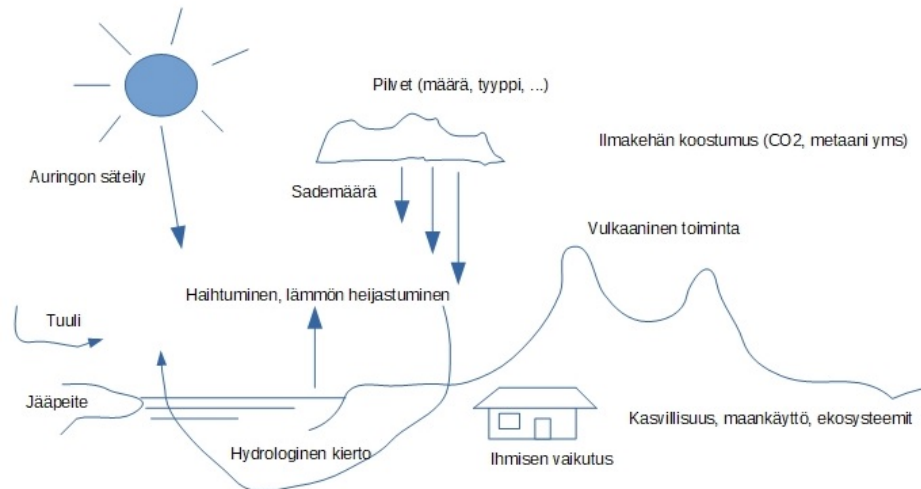
Erilaiset mittaustiedot ovat tyypillisiä ilmastonmuutoksen tutkimisessa käytettyjä suuraineistoja. Mittaustiedot voivat koostua esim. eri paikoista kerätyistä erilaisista säätiedoista kymmenien tai satojen vuosien ajalta.

Maapalloon ja ilmastoon liittyy paljon erilaista tietoa, jota pyritään keräämään ja analysoimaan, jotta ilmastonmuutokseen liittyviä asioita ja ilmiöitä voidaan paremmin ymmärtää. Tieto voi liittyä esim. auringon säteilyn muutoksiin, ilmakehän koostumukseen, ilmastoon, hydrologiseen kiertoon, pilvien esiintymiseen, tulivuorten aktiivisuuteen, jäätiköihin, haihtumiseen, maapallosta heijastuvaan lämpöön, tuulen ja sateen aiheuttamiin ilmiöihin, ihmisen aiheuttamiin muutoksiin, kasvillisuuden muutoksiin, veden kiertokulkuun, maankäyttöön ja muihin vastaaviin ilmiöihin. Näistä on olemassa paljon kerättyä dataa, jota on tyypillisesti kerätty tietyltä ajalta. Erilaisten sensorien fyysisen koon pieneneminen, teknologinen kehitys sekä yhteiskuntien lisääntynyt kiinnostus ymmärtää ilmastonmuutosta ovat myös johtaneet erilaisten kerättyjen data-aineistojen määrän kasvamiseen. Tyypillisesti data-analytiikassa pyritään luomaan erilaisia matemaattisia malleja tai luokittelemaan erilaisia asioita. Ilmastonmuutokseen liittyvässä data-analytiikassa haasteita tulee aineistojen suuresta määrästä sekä erityisesti siitä, ettei tutkittavia ilmiöitä välttämättä täysin ymmärretä. Tyypillisesti data-analytiikassa analysoidaan esim. yrityksen tietojärjestelmistä kerättyä tietoa, josta pyritään analysoimaan tarkkaan tiedettyjä asioita liittyen esim. kannattavuuteen ja liiketoiminnan ennustamiseen. Vastaavasti ilmastonmuutokseen liittyvissä aineistoissa ei välttämättä tiedetä, miten kerätty data kuvaa tutkittua ilmiötä ja millaisia monimutkaisia takaisinkytkentöjä ja riippuvuuksia ilmastoon liittyvällä datalla on erilaisiin asioihin (Faghmous et al. 2014, Guo et al. 2015). Jos tarkkaan rajattujen suuraineistojen analysointi liittyen hyvin tarkasti määriteltyihin ongelmiin (esim.

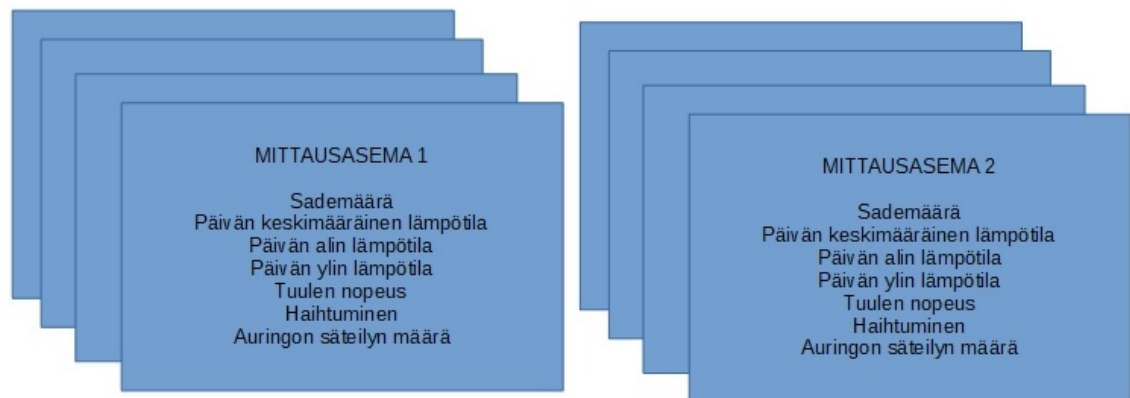
yrittäminen) on haasteellista, niin ilmastomuutokseen liittyvien suuraineistojen analysointi on todennäköisesti tästä vielä huomattavasti haasteellisempaa ja monimutkaisempaa.

NASA kerää aktiivisesti erilaisia suuraineistoja liittyen erilaisiin maapallon ilmiöihin ja tarjoaa esim. tarkkoja satelliittikuvia vapaasti käytettäväksi. NASA on kehittänyt tätä tarkoitusta varten omat tietojärjestelmät, joiden kautta kuka tahansa voi hyödyntää erilaisia suuraineistoja. Lisäksi NASAn keräämiä erilaisia suuraineistoja päivitetään automaattisesti. Suuraineistojen tallennuksessa on pyritty huomioimaan yhteensopivuus mm. aiempiin NASAn aineistoihin liittyen. Tyypillisesti suuraineistojen keräysprosessissa data muutetaan sopivaan yhdenmukaiseen muotoon, varmistetaan datan oikeellisuus, lisätään metadata sekä julkaistaan data vapaasti saataville. NASA tarjoaa myös erilaisista suuraineistoista tehtyjä analyysejä palveluina (Climate Analytics-as-a-Service). NASAn tuottamia suuraineistoja käyttää mm. IPCC (Schnase et al. 2015). Myös eri valtioilla ja muilla toimijoilla saattaa olla erilaisia vapaasti käytettävissä olevia suuraineistoja jotka soveltuvat ilmastomuutoksen tutkimiseen. Esimerkkinä valtiollisesta ilmastomuutokseen liittyvästä suuraineistosta on esim. yhdysvaltalainen Climate-Data.gov -sivusto, joka tarjoaa erilaisia suuraineistoja vapaasti käytettäväksi esim. tutkimustarkoituksiin (Climate-Data.gov 2019).

Myös Google tarjoaa erilaisia, ilmastoon ja maapalloon liittyviä suuraineistoja Google Earth Engine -palvelussaan. Palvelu on osa Googlen visiota liittyen maailmassa olevan datan järjestämiseen. Palvelua voi käyttää ilmaiseksi. Google Earth Engine käyttää osaltaan myös NASAn tarjoamia suuraineistoja (Google Earth Engine 2019). Science-lehden mukaan The British Geological Surveyin geologit suunnittelevat noin kolmen miljoonan fossiilin tietojen muuttamista digitaalseksi. Geologien mukaan nykyinen arkisto on kerätty 150 vuoden aikana mutta arkiston tietoja on hyvin vaikeata hyödyntää ja tavoitteena on luoda digitaalinen ”geobiodiversitettitietokanta” (Geobiodiversity Database) (Science 2019). Samalla myös Googlen Earth Engine pystyisi todennäköisesti hyödyntämään myös kyseistä tietokantaa.



Kuva 2. Ilmastoön liittyviä tyypillisiä suuraineistoja (mukailtu Faghmous et al. 2014).



Kuva 3. Suuraineisto voi koostua esim. mittausasemien keräämistä päivittäisistä säätiedoista (mukaillen Dietrich et al. 2019).

	Perinteinen data-analytiikka	Ilmastonmuutoksen data-analytiikka
Toimija	Yrityksen talousvastaava	Useita eri alojen ammattilaisia
Tilaaja	Yksittäinen yritys	Tyypillisesti valtio ja/tai suuryritykset
Tarkoitus	Yrityksen taloudellisten tunnuslukujen selvittäminen	Laajojen asioiden ymmärtäminen
Ennusteiden aikajänne	Yrityksen vuosineljännesraportointi	Kymmenistä vuosista eteenpäin
Aineiston aikajänne	Muutamia vuosia	Kymmenistä vuosista eteenpäin
Lähde	Yrityksen tietojärjestelmä	Useita erilaisia lähteitä
Tyypillinen data	Taloudellinen tunnusluku	Mittausdata
Prosessi	Talouden tunnusluvut tarkoin määriteltä	Paljon epävarmuutta
Loppukäyttäjät	Yrityksen johto	Tyypillisesti valtio ja/tai suuryritykset

	Perinteinen data-analytiikka	Ilmastomuutoksen data-analytiikka
Kaupallinen arvo	Lyhytaikaista arvoa yritykselle	Arvon määrittely hankalaa

Taulukko 2. Perinteisen data-analytiikan ja ilmastomuutokseen liittyvän datan vertailua (mukailtu soveltaen Faghmous et al. 2014).

Yllä olevat kuvat esittävät tyypillisiä ilmastomuutokseen liittyvässä suuraineistoissa käytettyjä mittaustietoja sekä perinteisen data-analytiikan ja erityisesti ilmastomuutokseen liittyvän data-analytiikan vertailuja hyvin karkealla tasolla.

Ilmastomuutokseen liittyvät suuraineistot saattavat sisältää myös paljon erilaista epävarmuutta. Mikäli säätietoja on kerätty sadan vuoden aikana manuaalisesti, voi esim. käytetyn mittalaitteen virhe olla jäänyt huomioimatta tai mittauksiin voisi sisältyä erilaisia mittausvirheitä. Vastaavasti yritysten tietojärjestelmien sisältämät tiedot ja niihin liittyvät asiat saatetaan tuntea paremmin.

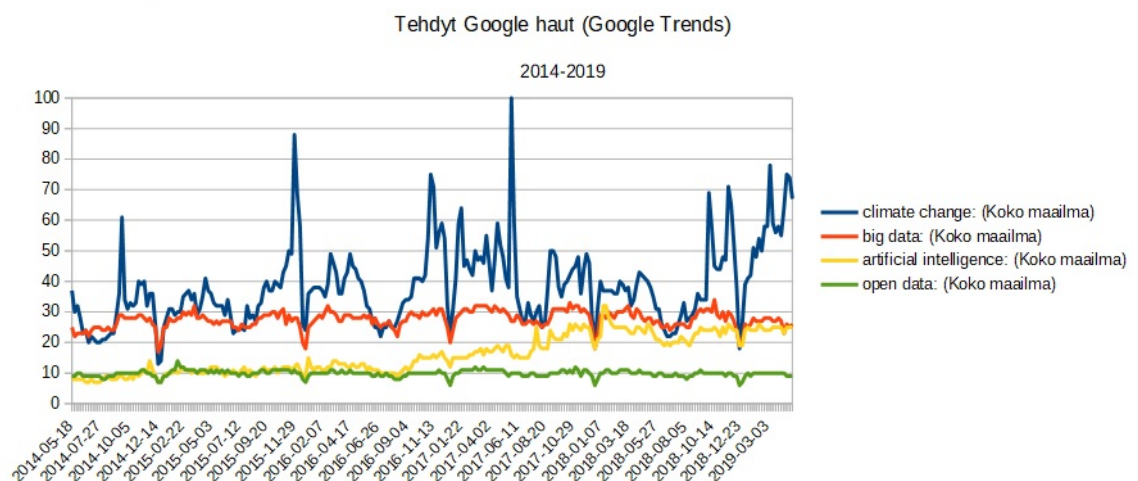
Tyypillinen ilmastomuutosta tutkiva ja suuraineistoja käyttävä tutkimus on esim. Die Erde -lehdessä julkaistu tutkimus, jossa on analysoitu ilmastomuutoksen vaikutusta Berliinin alueen säähän. Käytettävissä on ollut lämpötilatietoja vuodesta 1701 lähtien, mutta ensimmäisten 150 vuoden aikana mittaukset on tehty eri mittalaitteilla ja eri paikoissa, joten kaikki data ei ole kovin luotettavaa. Lisäksi Berliinin kaupungistuminen on aiheuttanut paikallista lämpenemistä, kun auringon lämpöä imeytyy rakennuksiin enemmän. Kaupungistumisen aiheuttama vaikutus lämpötiloissa on myös otettava huomioon. Lisäksi tutkimuksessa on tutkittu myös Berliinin alueen sademääriä. Tutkimuksessa on ensin analysoitu kerättyä dataa, josta on tehty mallinnus siitä, miten Berliinin alueen lämpötila ja sademäärä muuttuvat tulevaisuudessa sekä arvioitu, miten sama lämpötilanmuutos näkyisi globaalisti. Lisäksi myös tulevaisuuden sademääriä arvioitiin Saksan alueella. Lisäsimulaatioissa oli käytetty muiden tutkijoiden aiemmin tekemiä malleja. Tutkimuksen perusteella Berliinin alueen lämpötila nousee 3-3.5 astetta 21 vuosisadan loppuun asti sekä sateet lisääntyvät talvisin ja vähenevät kesäisin (Cubasch et al. 2011). Tämänäyttylinen tutkimus on hyvä esimerkki erilaisten suuraineistojen käyttämisestä ilmastomuutokseen liittyvässä tutkimuksessa. Tutkimuksessa ei selvitetty sitä, millainen vaikutus lämpötilan kasvamisella ja sademäärän muutoksilla on keskenään. Ilmaston lämpeneminen saattaa lisätä meriveden haihtumista, joka vaikuttaa sademääriin. Ilmastomuutokseen liittyvässä tutkimuksessa on myös tyypillistä se, että kahden asian (esim. Berliinin lämpötila ja sademäärä) keskinäistä vuorovaikutusta on hankala määritellä. Todennäköisesti käyttämällä useampia muita suuraineistoja, ilmaston toiminnasta saataisiin ehkä kattavampaa tietoa. Suuraineistojen sisältämän tiedon monimutkaisuus ja tulkitsemiseen liittyvät haasteet ovat tyypillisiä suuraineistoille.

3.5 Sosiaalinen media suuraineiston lähteenä

Mittaustietojen lisäksi myös erilaisista sosiaalisen median ja muiden vastaavien sähköisten lähteiden tuottamasta datasta voidaan tehdä analyyskejä, joita voidaan käyttää ilmastomuutokseen liittyvässä tutkimuksessa.

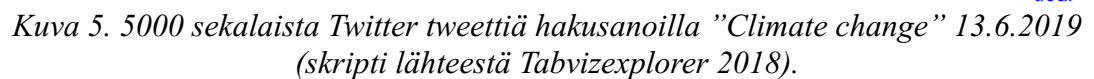
Myös sosiaalinen media tarjoaa erilaisia suuraineistoja joita voi kuka tahansa hyödyntää. Sosiaalisen median suuraineistoja ei välttämättä vielä osata hyödyntää ilmastonmuutokseen liittyvässä tutkimuksessa. Vuonna 2015 julkaistussa tutkimuksessa oli analysoitu ja luokiteltu Twitter -dataa ja tutkittu, missä (Yhdysvallat, Iso Britannia, Kanada ja Australia) alueilla ilmastonmuutoksen kieltäjät ovat (Jang et al. 2015). Vuonna 2019 julkaistuissa tutkimuksissa käsiteltiin kyllä sosiaalista mediaa ilmastonmuutoksen yhteydessä, mutta varsinaisesti suuraineistoja analysoivia tutkimuksia oli hyvin vähän. Tästä johtuen tässä kappaleessa käydään lyhyesti läpi, miten sosiaalisen median suuraineistoja voisi hyödyntää ilmastonmuutokseen liittyvässä tutkimuksessa. Esimerkkeinä on hyödynnetty Googlen hakukoneen (Google Trends) vapaasti saatavilla olevia aineistoja ja ilmaisen R Studio -analytiikkaohjelmiston ilmaisia analyysityökaluja. Esimerkkeinä on käytetty yksinkertaisuuden vuoksi Googlen lisäksi Twitter-palvelua ja myös muita sosiaalisen median (esim. Facebook) aineistoja on mahdollista samalla tavalla analysoida ilmastonmuutoksen tutkimuksen näkökulmasta.

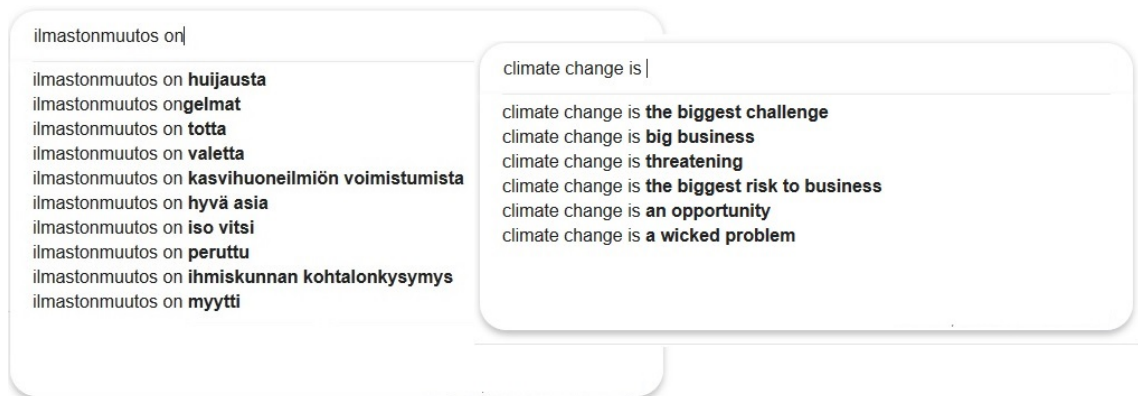
Google Trends -palvelun mukaan maailmalla ollaan haettu yhä enempi tietoa ilmastonmuutoksesta. Googella tehtyjen hakujen mukaan erityisesti ilmastonmuutoksesta on haettu yhä enemmän tietoa ja erityisesti vuoden 2019 alussa tiedonhaku on ollut vilkasta. Myös kiinnostus tekoälyyn on ollut kasvussa, mutta suuraineistoihin ja avoimeen dataan liittyvät haut ovat pysytelleet melko samoina vuosina 2014-2019. Vuonna 2017 Etelä-Euroopassa oli ennätyshelteet (Yle 2017a). Google Trends -palvelun raporttoimien hakujen mukaan ilmastonmuutokseen liittyvät haut ovat olleet ennätysmatalia joka jouluna 2014-2019 (Google Trends 2019), jolloin tyypillisesti länsimaissa keskitytään kuluttamiseen. Ehkä Google Trendsin aineiston perusteella joululahjojen jako pitäisi pohjoisella pallonpuoliskolla ajoittaa joulun sijasta kesään, jolloin ilmastonmuutostietoisuus olisi todennäköisesti huipussaan. Pystyakseli on Google Trendsin käyttämä oma suhteellinen arvo, joka skaalaa kuvaajan ymmärrettävämpään muotoon.



Kuva 4. Googlen käyttäjien hakuja (lukumäärät) ilmastonmuutokseen ja suuraineistoon liittyen 14.5.2019.

Tyypillisesti avointa dataa käyttävissä tutkimuksissa on hyödynnetty vapaasti saatavilla





Kuva 6. Googlen hakukoneen ehdotelmia sille, mitä ilmastonmuutos on hakujen perusteella 25.6.2019.

Tyypillisesti ilmastonmuutoksen liittyvässä tutkimuksessa käytetään erilaisia mittaustietoja. Tyypillinen mittaustieto voi olla useiden sääasemien lämpötilat esim. sadan vuoden aikana. Näiden lisäksi on mahdollista käyttää myös muita suuraineistoja, joista esimerkkinä on yllä olevassa kuvassa tehty Twitter-viestien analyysi sekä Googlen hakukoneen ehdotelmia. Twitter-analyysissä on käytetty twitteR-kirjasto R Studio -ohjelmassa sekä netistä löytyvää valmista ohjelmaskriptiä. Käytetty skripti haki Twitteristä viestejä (tweetit) jotka pyrittiin luokittelemaan (esim. ilo, suru, viha yms.) ja lopuksi luokitellut viestit visualisoitiin sanojen yleisyyden perusteella. Ilmastonmuutoksen tutkimuksen kannalta voi olla oleellista se, millaisia suuraineistoja analysoidaan. Yllä olevan Twitter-kuvan satunnaisena ajanhetkenä automaattisesti kerätyistä Twitter-viesteistä voi olla hyvin hankala saada järkevää analyysia vaikka viestit pystytään systemaattisesti analysoimaan. Vastaavasti ihmiset saattavat kirjoittaa rehellisemmin Googlen hakuihin asioista, josta oikeasti ollaan kiinnostuneita (Twitterin viestit näkyvät muille, Googlen haut vain käyttäjälle ja Googlelle). Sen sijaan esim. sääasemien lämpötila-arvot ja muut mitatut suureet voivat toimia paremmin analyysien lähteenä. Vastaavasti sata vuotta sitten kerätyistä lämpötila-arvoista ei välttämättä tiedetä, millä mittalaitteella ja millä tavalla lämpötila-arvot ovat kerätty kun taas Twitter-datan keräystapa ja keräyslogiikka on hyvin tiedossa. Lisäksi Twitterin toimintatapa tunnetaan yleisesti mutta ilmaston lämpötilojen vaihtelun kaikkia syy- ja seuraussuhteita ei välttämättä tunneta.

Helsingin Yliopistossa on tehty pro gradu -tutkielma, jonka yhtenä osana on analysoitu ilmastonmuutosta tekevien tutkimusorganisaatioiden sosiaalisen median viestintää. Pro gradu -tutkimuksen sosiaalisen median sisällön analysoinnin yhtenä tuloksena on se, ettei ilmastonmuutokseen liittyvä tutkimus juurikaan kiinnosta suurta yleisöä kuin lähinnä säähän liittyvien ilmiöiden osalta (Lehikoinen 2014). Voi olla, että kenties myöskään suuren yleisön luoma sosiaalisen median sisältö ei välttämättä liity suoraan ilmastonmuutokseen, vaikka tilanne toki voi olla nyt vuonna 2019 erilainen kuin vuonna 2014.

Sosiaalisen median datan käyttöön liittyen on varmaan syytä määritellä, että minkä tyyppistä dataa käytetään. Kännykän paikkatieto voi olla luotettavampaa kuin esim. kännykästä tehty sosiaalisen median päivitykset. Googlessa työskennellyt data-analyytikko Stephens-Davidowitz on analysoinut kirjassaan ihmisten sosiaalisen median päivityksiä (esim. Facebook) ja vastaavasti Googleen tehtyjä hakuja. Stephens-

Davidowitzin mukaan Googleen tehdyt haut todennäköisesti antavat realistisemman kuvan ihmisten käyttäytymisestä kuin esim. Facebookin päivitykset, joilla ehkä ennemminkin luodaan omaa positiivista imagoa (sivut 105-164, Stephens-Davidowitz 2017). Esimerkiksi Twitterissä käyttäjät saattavat tyypillisesti viestittää onnistuneesta lomasta (tiedon luotettavuus voi olla kyseenalainen) ja samaan aikaan hakea Googleasta tietoja esim. melanooman tunnusmerkeistä (tieto mahdollisesti luotettavampaa kuin Twitterin vastaava lomapäivitys).

On myös huomattava, että sosiaalisen median aineisto voi sisältää muutakin, kuin pelkkiä Twitter tai Facebook päivityksiä. Erilaiset ihmisten käyttämät laitteet voivat luoda erilaisia suuraineistoja, joita on mahdollista hyödyntää ilmastonmuutokseen liittyvässä tutkimuksessa. Vaikka vielä ei välttämättä tiedetä, miten tämänytyylisiä aineistoja voidaan tehokkaasti hyödyntää ilmastonmuutokseen liittyvässä tutkimuksessa, tämänytyypisten suuraineistojen olemassaolo kannattanee tiedostaa.

3.6 Muita kuin suuraineistojen käyttöön perustuvia ilmastonmuutostutkimuksia

Tämä diplomityö käsittelee pelkästään tutkimuksia, joissa on hyödynnetty erilaisia suuraineistoja. Ilmastonmuutosta tutkitaan kuitenkin monella muullakin tavalla, mutta aiheen rajaamiseksi nämä muut tutkimustyytit on jätetty tästä diplomityöstä pois. Tyypillinen ei-suuraineistoja käyttävä ilmastonmuutokseen liittyvä tutkimus on esim. asiantuntijahaastattelusta analysoimalla tehty tutkimus.

Suuraineistojen käyttäminen ilmastonmuutokseen liittyvässä tutkimuksessa on yleistä erityisesti biologian ja muiden luonnontieteiden alalla, jossa suuraineistoja on yleensä valmiiksi saatavilla. Osa ilmastonmuutokseen liittyvistä tutkimuksista on toteutettu erilaisin haastatteluin, kuten esim. tutkimus ilmastonmuutoksen vaikutuksesta Australian rakennusteollisuuteen (Hurlimann et al. 2019) tai ilmastonmuutoksen vaikutusta Ison Britannian tieverkostoon (Wang et al. 2019). Erilaisista yhteiskunnallisista asioista voi olla hyvin hankala saada numeeriseen muotoon muutettua yksikäsitteistä tietoa. Myös erilaisia puhtaasti kirjallisuustutkimukseen perustuvia ilmastonmuutostutkimuksia on olemassa. Näistä esimerkkinä on kirjallisuustutkimus, jossa on kerätty eri lähteistä (aiempia tutkimuksia) ilmastonmuutoksen aiheuttamia uhkia hiihtoturismille (Steiger et al. 2019). Lisäksi on olemassa tutkimuksia, joissa on yhdistetty erilaisia todennäköisesti suuraineistoja käyttäneitä tutkimuksia. Näistä esimerkkinä on tutkimus otsonikadon ja ilmastonmuutoksen yhteisvaikutuksesta meren ekosysteemiin (Williamson et al. 2019).

Fortune-lehdessä julkaistun artikkelin mukaan erilaisten sensorien tuottaman datan määrä maailmassa lisääntyy. Ilmastonmuutokseen liittyvässä tutkimuksessa haasteena on Fortuneen mukaan se, että erilaisista suuraineistoista tehdyistä analyyseistä vain osa on oikeita ja käyttökelpoisia. Tämä johtuu Fortuneen mukaan siitä, että erilaisten ilmastoon liittyvien ilmiöiden keskinäisiä vuorovaikutuksia (syy- ja seuraussuhteita) ei välttämättä vielä ymmärretä ja analyysien johtopäätökset voivat osoittautua vääriksi (Fortune 2016).

3.7 Aiempia diplomitöitä ja muita julkaisuja aiheesta Suomessa

Varsinaisia kirjallisuuskatsauksia tai tilannearviota suuraineistojen käyttämisestä ilmastonmuutokseen liittyen ei ole juurikaan tehty. Olemassa olevat lopputyöt käsittelevät lähinnä suuraineistojen käyttöä yritysten liiketoiminnassa tai tutkittavien asioiden selvittämistä käyttäen erilaisia suuraineistoja. Aiemmin on tehty lopputöitä liittyen ilmastonmuutoksen tutkimiseen ja osassa tämän tyyllisiä lopputöitä on hyödynnetty erilaisia suuraineistoja. Erilaisten laajojen aineistojen käyttö tutkimuksissa ei ole sinänsä uusi asia ja satojen vuosien aikana on tehty lukuisia tutkimuksia erilaisia suuria aineistoja keräämällä ja tekemällä näistä erilaisia johtopäätöksiä. Tyypillisesti aineistoihin on voinut olla rajoitettu pääsy eivätkä aineistot ole välttämättä olleet digitaalisessa muodossa.

3.7.1 Aiempia suuraineistoihin liittyviä tutkimuksia

Suuraineistojen hyödyntämiseen liittyen on kirjoitettu lukuisia artikkeleita ja tutkimuksia. Tampereen Teknillisessä Yliopistossa on tehty diplomityö, missä on tutkittu suuraineistojen maturiteettimalleja. Diplomityössä on vertailtu eri tapoja hyödyntää suuraineistoja yrityksen liiketoiminnassa ja erityisesti sitä, millä tavoin hyödyntämällä suuraineistoista saadaan suurin hyöty lisäarvon tuottajana yrityksen liiketoimintaprosesseihin. Tyypillisesti yrityksen suuraineistojen hyödyntämisprosessi on jatkuva prosessi, missä yrityksen toimintaa (esim. suuraineistojen analysointia) pyritään jatkuvasti parantamaan ja prosessia kehittämään. Suuraineistojen hyödyntämistavat eivät tyypillisesti ole standardisoituja eivätkä välttämättä hyvin dokumentoituja (Braun 2015). Suuraineistojen hyödyntäminen erilaisten organisaatioiden toiminnassa on ehkä hieman haparoivaa ja mahdollisesti myös liiketoimintakohtaista. Tuskin myöskään on yksiselitteistä määritelmää sille, miten suuraineistoja hyödynnetään optimaalisesti erityisesti ilmastonmuutokseen liittyvässä tutkimuksessa.

Toisessa Tampereen Teknillisessä Yliopistossa tehdyssä diplomityössä on tutkittu sitä, millä tavoin datan ja data-analytiikan tulosten myynnillä voidaan luoda liiketoimintaa. Tästä liiketoiminnasta käytetään termiä "data monetization", jolle ei ole vielä suomenkielistä vastinetta. Tutkimuksen mukaan arvokkaan datan tunnistaminen ja uuden liiketoiminnan luominen data-analyysien tuotosten perusteella on keskeistä. Kuitenkin "data monetization" on terminä hyvin monimutkainen (Fred 2018). Ehkä myös tämä tutkimus antaa viitteitä siitä, että suuraineistojen laajempi kaupallinen hyödyntäminen ei ole vielä tarkkaan hahmottunut ja ala on vasta kehittymässä. On myös mahdollista, että tämän tyylinen dataliiketoiminta on myös yritysten liikesalaisuus, jolloin aiheesta ei välttämättä ole julkisesti saatavilla tarkkoja tietoja. Toisaalta Lappeenrannan Teknillisessä Yliopistossa tehdyssä väitöskirjassa todetaan, etteivät yritykset edes välttämättä halua antaa omia data-aineistojaan muiden käyttöön, koska yritykset haluaisivat itse vaikuttaa siihen, miten ja millaisia kaupallisia palveluita datasta luodaan (Herala 2018). Ilmastonmuutokseen liittyvät tutkimukset hyödyntävät tyypillisesti erilaisten valtiollisten toimijoiden vapaasti saatavilla olevaa dataa, ja ehkä tutkimuksia rajoittaa myös se, ettei käytössä välttämättä ole yritysten keräämiä

suuraineistoja. Tämä voi johtaa siihen, että ilmastonmuutokseen liittyvät data-analyysit koskevat tyypillisesti ei-kaupallisia asioita. Toisaalta Tampereen Yliopistossa tehdyn väitöskirjan mukaan yritykset pyrkivät vaikuttamaan ilmastonmuutokseen liittyvään politiikkaan siten, että poliittiset ratkaisut tukevat yritysten kykyä säilyttää oma liiketoiminta tulevaisuudessakin (Kentala-Lehtonen 2019).

Tampereen Teknillisessä Yliopistossa on tehty diplomityö, missä on tutkittu avoimen datan lähteiden hyödyntämisen haasteita big data analytiikassa. Tutkimuksessa tarkasteltiin loppukäyttäjän näkökulmasta 16 avoimen datalähteen (mm. säätietoja) laatua, kokonaisuutta sekä saatavuutta. Diplomityön mukaan avointen datalähteiden käyttämiseen liittyvät haasteet ovat erilaisten tiedostomuotojen eroavaisuudet, metadatatassa olevat puutteet, havaintotarkkuuksien erot, maantieteelliset rajoitteet, heikko arkkitehtuurikuvaus sekä rajapinnat, datan laadun eroavaisuudet ja heikko löydettävyyttä sekä saatavuus. Diplomityön mukaan ratkaisuna haasteisiin olisi yhtenäisen määritelmän luominen avoimelle datalle sekä sille, miten avointa dataa tulisi tarjota (Heinonen 2016). Osa yrityksistä todennäköisesti tekee erilaisia ilmastonmuutokseen liittyviä suunnitelmia joten tämän mainitun diplomityön mainitsemat haasteet ovat todennäköisesti samoja ilmastonmuutokseen liittyvässä avoimen datan käyttämisessä. Todennäköisesti myös valtiolliset toimijat ja muut vastaavat organisaatiot kamppailevat samantyylisten ongelmien kanssa.

Suuraineistoista ja niiden käyttämisestä löytyy näiden muutaman referoidun tutkimuksen lisäksi paljon muitakin tutkimuksia. Kun suuraineistojen hyödyntämiseen lisätään ilmastonmuutos ja ilmastonmuutokseen liittyvien haasteiden ratkaisu, aihe muuttuu hyvin monimutkaiseksi. Pelkästään suuraineistojen käyttämisen ja ilmastonmuutoksen ongelmien kuvaaminen voi olla hyvin haastavaa, eikä välttämättä käytössä ole tietoa siitä, mitä lopputuloksena pitäisi olla. Lisäksi ilmastonmuutoksen vaikutuksia ei välttämättä tunneta tarkasti.

Ilmastonmuutos ja suuraineistot saatetaan helposti mieltää erillisiksi ongelmiksi tai haasteiksi. Ehkä kuitenkin loogisempi lähtökohta on nähdä suuraineistojen käyttäminen yhtenä työkaluna ilmastonmuutokseen liittyvässä tutkimuksessa. Suuraineistojen luomat ongelmat ja haasteet ovat ehkä vain tämän tutkimustavan (suuraineistojen analysoinnin) haasteita. Ilmastonmuutokseen ei välttämättä vielä liity selkeästi hahmotettavaa liiketoimintaa, jolloin ilmastonmuutokseen liittyvät suuraineistojen analysoinnit saattavat liittyä enempi ei-kaupalliseen tutkimukseen.

3.7.2 Aiempia ilmastonmuutosta ja suuraineistoja yhdisteleviä tutkimuksia

Tyypillinen ilmastonmuutosta ja suuraineistoa käsittelevä opinnäytetyö on esim. Turun Yliopistossa julkaista väitöskirja, jossa on tutkittu ilmastonmuutoksen vaikutusta Itämeren silakka- ja planktonkantoihin. Vaikka tutkimuksessa ei suoraan puhuta suuraineistosta, käytettävissä on kuitenkin ollut planktoniin liittyviä havaintoja (kerättyjä näytteitä planktonin esiintymisestä) Itämerellä ajalla 1967-1984 ja 1991-2013 (vuosien 1984-1991 data puuttui datan kerääjäyrityksen vaihtumisesta johtuvista syistä). Lisäksi käytössä oli veden laatuun ja lämpötiloihin liittyvää aineistoa sekä silakkakantoihin liittyvää tietoa. Tutkimuksessa on yhdistelty useita aineistoja mm.

Ilmatieteenlaitoksen keräämiä säähavaintoja sekä tekijän itsensä tekemiä havaintoja. Käytetyt data-aineistot on yhdistetty ja näitä on analysoitu tilastollisin menetelmin. Tutkimuksen mukaan ilmastonmuutos on vähentänyt eläinplanktonin määrää Saaristomerellä sekä aikaistanut planktonin vuodenaikaista esiintymistä. Lisäksi silakkakannan kasvu sekä meriveden makeutuminen ja lämpeneminen ovat yhteydessä silakan rasvapitoisuuden kasvuun (Mäkinen 2019). Vaikka väitöskirjassa ei esiinny big data- tai suuraineisto-termiä, on tämäntyylinen tutkimus tyypillinen esimerkki ilmastonmuutokseen liittyvästä erilaisten osittain vapaastikin saatavilla olevien suuraineistojen käytöstä. Hieman samantyyppinen väitöskirja on tehty Oulun Yliopistossa liittyen Pohjois-Fennoskandian kasvillisuusmuutoksiin. Väitöskirjassa on tutkittu vuosina 1957-1981 kerättyjä kasvinäytteitä, joita on verrattu vuosina 2013-2014 kerättyihin kasvinäytteisiin. Lisäksi tutkimuksessa on käytetty erilaisia valtion ylläpitämiä avoimen datan aineistoja kuten alueen lämpötiloja (Maliniemi 2018). Tyypillisesti erilaisissa tämän tyyllisissä tutkimuksissa on käytetty erilaisia suuria aineistoja sekä yhdistelty että myös analysoitu sekä mallinnettu tutkittavia asioita. Kuitenkin vanhemmissa tutkimuksissa saattaa olla käytetty pienempiä datamääriä kuin mitä ehkä uudemmissa tutkimuksissa on käytetty. Lisäksi aiemmissa tutkimuksissa tekijät ovat olleet usein vain tietyn melko tarkkaan rajatun alan (esim. biologia) liittyviä henkilöitä joilla on valmiudet tilastolliseen analyysiin. Mikäli tutkimusten aineistoja, maantieteellistä kattavuutta sekä käytettävissä olevien asiantuntijoiden erikoisalojen määriä lisättäisiin, niin todennäköisesti saataisiin kattavampaa tietoa ilmastonmuutoksen vaikutuksista. Todennäköisesti kuitenkin nämä kaksi esimerkkinä käytettyä väitöskirjaa toimivat suuremman kokonaisuuden osina, kun ilmastonmuutoksen kokonaisvaikutuksia tutkitaan.

Valtioneuvosto on julkaissut ilmastonmuutoksen heijastevaikutuksista Suomeen artikkelin, jossa tarkastellaan ilmastonmuutoksen vaikutuksia yhteiskunnan eri toimintoihin erilaisten monimutkaisten vuorovaikutussuhteiden kautta (Hildén et al. 2016). Vaikka raportissa ei erityisesti käsitellä erilaisten suuraineistojen hyödyntämistä, ovat monet yhteiskunnan järjestelmien vuorovaikutussuhteet ja ilmastonmuutoksen vaikutukset todennäköisesti hyvin hankalasti hahmotettavissa. Tällöin päätöksenteossa todennäköisesti tarvittaisiin erilaisten suuraineistojen analysointia päätöksenteon tueksi sekä erilaisten yhteiskunnan järjestelmien riippuvuuksien hahmottamiseksi.

Tampereen Teknillisessä Yliopistossa on tehty diplomityö liittyen maanteiden kuivatukseen ja kuivatusjärjestelmiin liittyvään tiedonhallinnan kehitystarpeisiin ilmaston muuttuessa. Diplomityön mukaan erityisesti lisääntyvät sademäärät saattavat aiheuttaa lisääntyneen maanteiden kuivatustarpeen ja nykyisellään käytössä on erillisiä tietojärjestelmiä. Ratkaisuksi ehdotetaan tietojärjestelmien yhdistämistä muihin tietojärjestelmiin (mm. hydrologisia havaintoja tarjoaviin tietojärjestelmiin) (Toivonen 2016). Diplomityö kuvastaa hyvin suurten datamassojen analysointia ja yhdistämistä, sekä lopputulosten käyttämistä tietyn tehtävän (liikennejärjestelmän huollon) optimoinnissa. Lienee todennäköistä, että vastaavanlaisia ratkaisuja tullaan tarvitsemaan muissakin vastaavissa järjestelmissä eri toimialoilla.

Aalto Yliopiston väitöskirjassa käsitellään paikkatietoanalyysin käyttämistä laajan avoimen datan kanssa ja lopputuloksena saatua tietoa hyödynnetään vesivarojen hallinnassa. Analysoinnissa suositellaan käytettävän useampaa analysointitapaa, selkeää

raportointia sekä tietolähteiden luokittelemista (Salmivaara 2015). Väitöskirja on hyvä esimerkki tilanteesta, jossa käytettävissä on runsaasti erilaista dataa, jota analysoimalla saadaan tukea päätöksenteolle. Oletettavasti käytetyn datan ja analyysien luokittelemisella voidaan saada lisäarvoa päätöksentekoon, koska luokittelemalla voidaan esim. varmistua käytetyn datan tai johtopäätösten luotettavuudesta.

Tampereen Yliopiston julkaisemassa väitöskirjassa on analysoitu Tampereen linja-autoliikenteestä kerättyä dataa, jota analysoimalla on voitu tarkastella liikenteen sujuvuutta eri tilanteissa sekä eri liikennöintilinjoilla (Syrjärinne 2016). Todennäköisesti tällaiseen analyysiin voitaisiin yhdistää myös muualta kerättyä tai saatavilla oleva dataa (esim. avoin data) sekä myös simuloida erilaisia mahdollisia ilmastomuutoksen (esim. rankkasateet tai muut äärisääilmiöt) vaikutuksia.

Tampereen Yliopiston johtamiskorkeakoulussa tehdyssä pro gradu -tutkielmassa on analysoitu Eduskunnassa käytävää ilmasto- ja energiastategiasta käytyä lähetekeskustelua marraskuussa 2016. Keskustelussa nousi esiin tekninen (toimenpiteitä, teknisiä ratkaisuja yms.), ympäristö (vaikutukset ympäristöön yms.), eristävä (reagointi ympäristöongelmiin aiheuttaa uusia suuria ongelmia) sekä valoisa (ongelmat ovat voitettavissa) näkökulma (Mentula 2018). Nykyinen globaali talousjärjestelmä tuottaa hyvin paljon dataa jota analysoimalla todennäköisesti saataisiin lisätietoa ilmastomuutoksen vaikutuksesta yhteiskuntaan. Pro gradu -tutkielmassa ei oltu tehty mitään visualisointeja esim. keskustelujen avainsanoista eikä muutakaan vastaavantyylistä visuaalista esitystä (esim. sanapilveä). Tyypillisesti ilmastomuutokseen liittyvä data mielletään helposti jonkinlaiseksi mittaustietoaaineistoksi, mutta aihetta voidaan tutkia myös erilaisia tekstejä ja sosiaalisen median dataa analysoimalla.

Hieman samantyylinen tekstianalyysi on tehty Helsingin yliopiston matemaattis-luonnontieteellisessä tiedekunnassa koskien sitä, miten lehdissä kirjoitellaan maaseudusta ilmastomuutokseen liittyen. Tutkimusaineisto koostui Helsingin Sanomien artikkeleista vuosina 1990-2012, painottuen vuosiin 2005-2011. Uutisaineistoa analysoimalla kirjoitukset luokiteltiin viiteen eri kategoriaan (diskurssiin): maaseudun mahdollisuudet, maaseutu saastuttajana, maaseudun puolustus, maaseutu ekologisena ja maaseudun haavoittuvuus (Lavonen 2013).

4 TULOKSET

Ilmastonmuutoksesta on tehty paljon tutkimusta viimeisen sadan vuoden aikana. Tyypillisesti on tutkittu esim. erilaisten elinympäristöjen muutoksia sekä kasvien ja eläinten leviämistä. Aiemmin käytössä on ollut lähinnä ei-digitaalista dataa kuten erilaisia muistiinpanoja, fyysisiä kasvi- ja eläinnäytteitä ja muuta vastaavaa. Nykyään erilaisten sensorien ja tietoverkkojen yleistymisen on kuitenkin tuonut rinnalle hyvin laajoja ja usein myös vapaasti saatavilla olevia aineistoja. Erilaisia mittaustietoja voidaan hyödyntää ja tietokoneiden laskentatehon kasvaminen on myös mahdollistanut yhä monimutkaisempien analyysien ja visualisointien tekemisen. Tyypillisesti ilmastonmuutokseen liittyviä suuraineistoja keräävät erilaiset valtiolliset ja ei-kaupalliset toimijat. Suuraineistoja on myös laitettu yhä useammin internetiin vapaasti saataville.

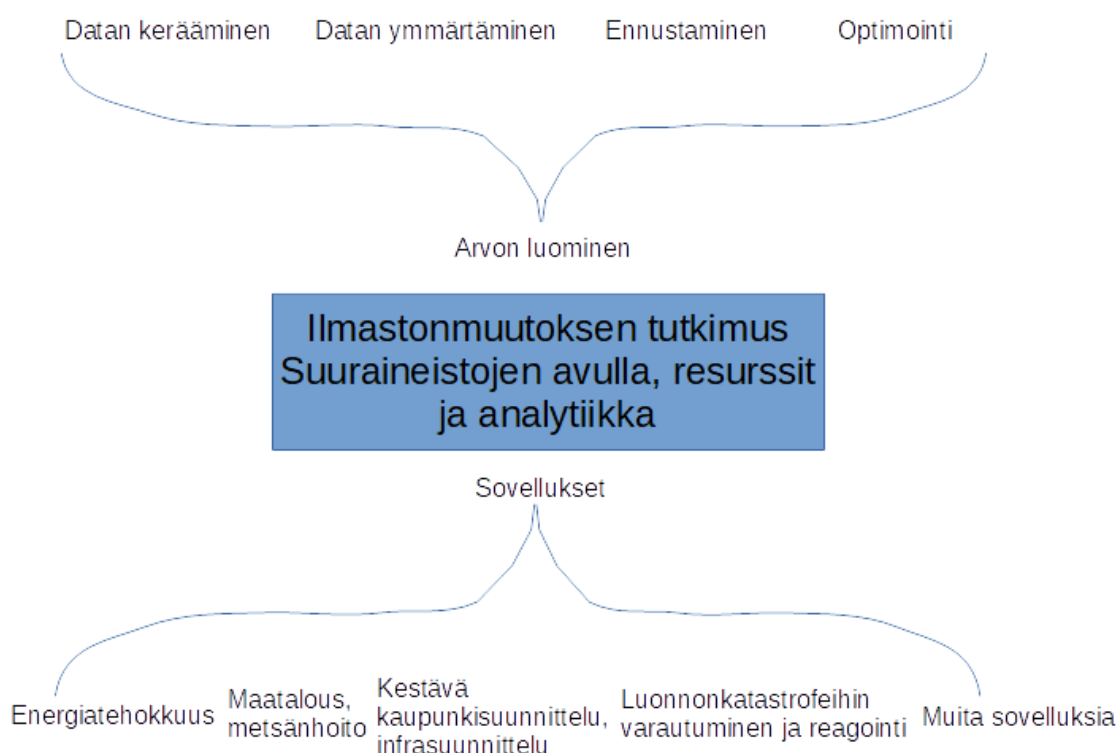
Tässä kappaleessa käydään läpi vuonna 2019 julkaistuja tutkimuksia, joissa on hyödynnetty erilaisia suuraineistoja ilmastonmuutokseen liittyvässä tutkimuksessa. Tutkimukset ovat jaoteltu yleisiin (suuraineistojen datan käyttäminen yms.), riskienhallintaan (erilaisiin uhkiin varautuminen) sekä aluespesifisiin (esim. tiettyä maantieteellistä aluetta koskeviin) tutkimuksiin. Jako on keinotekoinen ja tutkimukset voitaisiin jaotella muullakin tavoin. Tarkoituksena on kuitenkin saada karkea käsitys siitä, miten tutkimukset ovat karkeasti ottaen jakaantuneet.

4.1 Tutkimuksia liittyen käytettäviin data-aineistoihin

Tässä kappaleessa käsitellään tutkimuksia, joissa on käsitelty erilaisten suuraineistojen käyttämistä ilmastonmuutokseen liittyen. Pääpaino on enempi käytettävissä olevissa suuraineistojen ominaisuuksissa.

Big Data and Cognitive Computing -julkaisussa on käsitelty yleisellä tasolla erityisesti suuraineistojen käyttämistä ilmastonmuutokseen liittyvässä tutkimuksessa. Tutkimus käy läpi viime vuosina tehtyjä tutkimuksia ja pyrkii hahmottelemaan näistä yhteenvetoa. Tutkimuksen mukaan tyypillisesti tarvitaan havaintoja (suuraineistoja), havaintojen ymmärtämistä, ennustamista sekä järjestelmän optimointia. Erilaisia data-analytiikan työkaluja käyttämällä voidaan luoda lisäarvoa erilaisilla toimijoille. Näitä toimijoita ovat tyypillisesti älykkääseen energiaan ja energiansäätöön, maatalouteen ja metsänhoitoon, kestävään kaupunkisuunnitteluun, luonnonkatastrofien ja erilaisten tautien leviämiseen sekä muihin vastaaviin liittyvät toimijat. Data-analytiikkaan liittyvässä matematiikassa on tarjolla erilaisia työkaluja liittyen esim. datan luokitteluun ja ennustamiseen. Tyypillinen suuraineistoja käsittelevä ilmastonmuutostutkimus jakaantuu datan keräämiseen, datan ymmärtämiseen, ennusteiden tekemiseen ja erilaisiin optimointeihin liittyen luotuihin malleihin sekä tarkasteltavaan ongelmaan. Tämä prosessi luo datasta lisäarvoa ja erilaisia malleja voidaan käyttää tehtäessä

tutkimusta, joka luo lisäarvoa tutkimuksen loppukäyttäjille. Tyypillisesti suuraineistojen analysointiin liittyy runsaasti erilaisia data-analytiikkaa, koneoppimista, luokittelua ja muuta raskasta laskentatehoa vaativaa operaatiota (Hassani et al. 2019). Data-analytiikkaa ja suuraineistoja käytetään hyvin monessa sovelluksessa ja hyvin moneen tarkoitukseen. Ilmastomuutoksen luonne monimutkaisena ilmiönä johtaa kuitenkin siihen, että muuntuylliset (muut, kuin suuraineistojen analysointi) lähestymistavat voivat olla haasteellisia. Lisäksi ilmastomuutoksen monimaiset vaikutukset voivat heijastua muuallekin kuin esim. äärisäiden ennustamiseen. Suojelupoliisin Juhlavuosikirja 2018 - julkaisussa mm. ilmastomuutos ja luonnonvarojen mahdollinen resurssipula on lueteltu Suomen turvallisuutta uhkaaviksi tekijöiksi (Suojelupoliisi 2018, sivu 14). Vastaavasti yrityksillä voi olla tarpeen tukeutua erilaisiin ilmastomuutosta käsitteleviin suuraineistoanalyysihin omien toimintojensa suunnittelussa. Alla olevassa kuvassa on esitetty tyypillinen ilmastomuutosta käsittelevän tutkimuksen rakenne. On myös huomattava, että suuraineistoja käyttävissä tutkimuksissa voi olla käytössä muitakin tutkimustapoja (suuraineistojen analyysien rinnalla, esim. asiantuntijoiden haastatteluja datan ymmärtämiseen liittyen) ja alla oleva kuva esittää lähinnä perinteistä suuraineistoa hyödyntävää teknistä tutkimusta. Ilmastomuutokseen liittyvät sosiaaliset, taloudelliset ja muut vastaavat näkökulmat voivat myös vaatia muitakin lähestymistapoja tutkimuskysymyksiin (muita, kuin data-analytiikkaa ja matematiikkaa).



Kuva 7. Tyypillisen suuraineistoja käyttävän ilmastomuutostutkimuksen rakenne (Hassani et al. 2019).

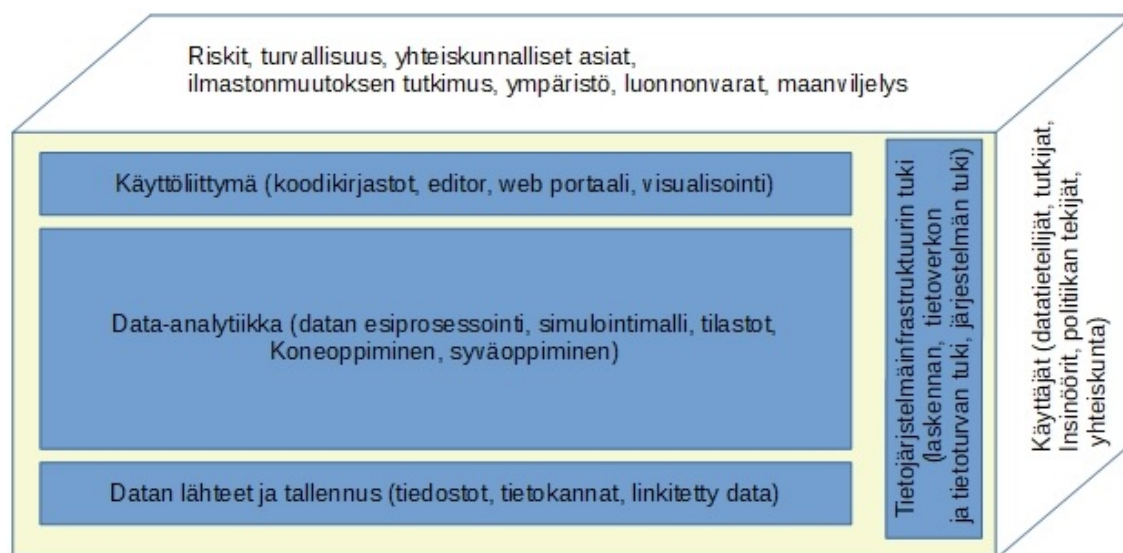
Edellämainitun Big Data and Cognitive Computing -julkaisussa on myös käyty läpi sitä, miten erilaisia suuraineistoja on viime vuosina hyödynnetty ilmastomuutokseen liittyvässä tutkimuksessa. Tutkimuksen mukaan suuraineistoja on viime vuosina hyödynnetty ilmastomuutokseen liittyvässä tutkimuksessa seuraavasti (Hassani et al. 2019):

- Energiatehokkuus: olemassa olevien resurssien käytön optimointi, älykkäät energiansiirtoverkot, energiankulutuksen arviointi, rakennuksiin liittyvät suuraineistot sekä rakennusten energiatehokkuuden optimointi, vesivoiman käyttöön liittyvien suuraineistojen hyödyntäminen
- Maatalous, metsänhoito: älykäs maanviljelys, maatalouden optimointi, ilmastonmuutosten vaikutusten huomiointi maa- ja metsätaloudessa, metsävarojen hallinta, maankäyttöön liittyvien suuraineistojen hyödyntäminen, erilaisten suuraineistojen yhdistäminen optimaaliseen maan ja metsänhoitoon liittyen, kestävä kehitys
- Kestävä kaupunkisuunnittelu, infrasuunnittelu: suuraineistojen hyödyntäminen kaupunkien ilmastonmuutokseen varautumisessa, liikennejärjestelmien suunnittelu, kaupunkien verkottuneiden laitteiden (IoT-laitteet) tuottaman datan hyödyntäminen, kaupunkien ilmastovaikutusten selvittäminen
- Luonnonkatastrofeihin varautuminen ja reagointi: ilmastonmuutoksen ja erilaisten luonnonkatastrofien vuorovaikutus, suuraineistojen hyödyntäminen tulvien, tulivuorenpurkausten, maanjäristysten ja muiden vastaavien ennakkoinnissa, sääilmiöiden mallintaminen, joukkoistamisen (crowdsourcing) ja miehittämättömien sensorien tuottaman datan hyödyntäminen luonnonkatastrofeissa, sosiaalisen median ja uutisten analysointi liittyen luonnonkatastrofeista toipumiseen, luonnonkatastrofien tuhojen arviointi, kulkutautien leviämisen mallintaminen ja ennakointi
- Muita sovelluksia: Sosio-ekonomisen datan analysointi kestävän kehityksen hengessä, suuraineistojen hyödyntäminen tuotteiden elinkaarenhallinnassa sekä valmistuksessa, suuraineistojen hyödyntäminen laitteiden hävittämiseen ja kierrätykseen liittyen

International Journal of Digital Earth -julkaisun tutkimuksen mukaan maapallolta kerätyllä suuraineisto datalla on tyypillisesti seuraavanlaisia piirteitä. Datan tulokset ovat ainutkertaisia eikä niitä voi toistaa. Kerättyyn dataan liittyy erilaisia epävarmuuksia, jotka vaikeuttavat datan tulkintaa. Data on tyypillisesti moniulotteista (monesta lähteestä eri tavoin ja eri aikoina kerättyä), koska esim. maapallon ekosysteemit, ilmasto ja muut vastaavat ovat kytkeytyneet toisiinsa monin eri tavoin. Näistä edellä mainituista johtuen tämäntyylisten suuraineistojen tulkinta vaatii runsaasti laskentatehoa ja tyypillisesti tarvittava matematiikka on hyvin monimutkaista. Vastaavasti suuraineistojen käsittelyä voidaan helpottaa kiinnittämällä huomiota esim. datan oikeellisuuteen, data-analysoinnin laskennan tehokkuuteen ja datan skaalautuvuuteen. Tyypillisesti julkisten lähteiden suuraineistoja on saatettu esikäsittellä ensin (esim. datan luettavuuteen liittyen) jonka jälkeen aineistot ovat saatavilla verkosta. Tutkimuksessa viitattiin sosiaalisen median käyttämiseen yhtenä datalähteenä, mutta tutkimuksessa ei kovin tarkasti yksilöity, miten tämä käytännössä tapahtuisi. Tyypillinen haaste suuraineiston analysoinnissa on tutkimuksen mukaan suuraineiston sisältämän informaation muuttaminen ymmärrettävään muotoon. Tutkimuksessa käytetään maapallosta kerätyistä suuraineistoista termiä ”Earth Observation Data” ja datan teknisen analyysin rinnalla tarvitaan aina laajempaa ymmärrystä datan keräämisestä ja datan sisällöstä. Haasteina suuraineistojen hyödyntämisessä on usein myös teknologia, datamassojen hallinta (access), analysoinnin (data processing) haasteet sekä analyysien tulosten jakaminen muille kiinnostuneille (Sudmanns et al. 2019). Vapaasti verkossa olevia suuraineistoja voi tarvittaessa käyttää erilaiset tutkimusryhmät, jotka kukin

tutkivat omia erikoisalueitaan. Tällöin yksittäinen suuraineisto voi hyödyttää useita eri tutkimusryhmiä eri puolilla maailmaa ja eri tutkimusryhmissä. Tällöin myös mahdolliset suuraineiston virheet monistuvat helposti kaikille käyttäjille. Toisaalta mikäli eri tutkimusryhmät tekevät tehokasta yhteistyötä, voidaan suuraineiston datasta tehtyjä analyysejä yhdistellä ja saada jälleen uutta tietoa. Yksi mahdollinen tulevaisuuden kehityssuunta voisi olla kaupallisten analyyssipalvelujen ilmaantuminen. Kaupalliset toimijat voisivat hyödyntää vapaasti saatavilla olevia suuraineistoja erilaisten analyysien tekemiseen. Todennäköisesti tämentyyllisiä markkinoita voisi olla esim. vakuutuksiin ja muuhun riskienhallintaan liittyen. Mahdollisesti tämän tyylisten kaupallisten sovellusten kehittäminen voi olla kallista ja mahdollinen asiakaskunta voi rajoittua yrityksiin, joiden liiketoiminta ei perustu lyhytaikaiselle ja ilmastosta riippumattomalle toiminnalle (esim. energiayhtiöt voisivat olla mahdollinen kohderyhmä).

Big Earth Data -julkaisussa on käsitelty maapallolta kerättyjä ja tutkimuksissa käytettyjä suuraineistoja (Big Earth Data). Tutkimuksessa on käyty läpi nykyisiä trendejä ja alla olevassa kuvassa on hahmoteltuna tyypillinen systeemiarkkitehtuuri tutkimuksiin, joissa suuraineistoja on käytetty ilmastomuutokseen liittyvässä tutkimuksessa. Tyypillisesti suuraineistojen käyttöön liittyy erilaisiin ryhmiin (esim. maatalous, luonnonvarojen käyttäminen) liittyviä tutkimuskysymyksiä. Aineistojen analysointiin liittyy tyypillisesti erilaisia teknisiä toimenpiteitä (esim. data-analyysijä, mallintamista, koneoppimista sekä näiden tukitoimintoja). Tutkimuksia tekemässä tuloksia analysoimassa on tyypillisesti eri alojen ihmisiä (esim. datatieteilijöitä, luonnontieteen tutkijoita, yhteiskunnan jäseniä) (Yang et al. 2019).



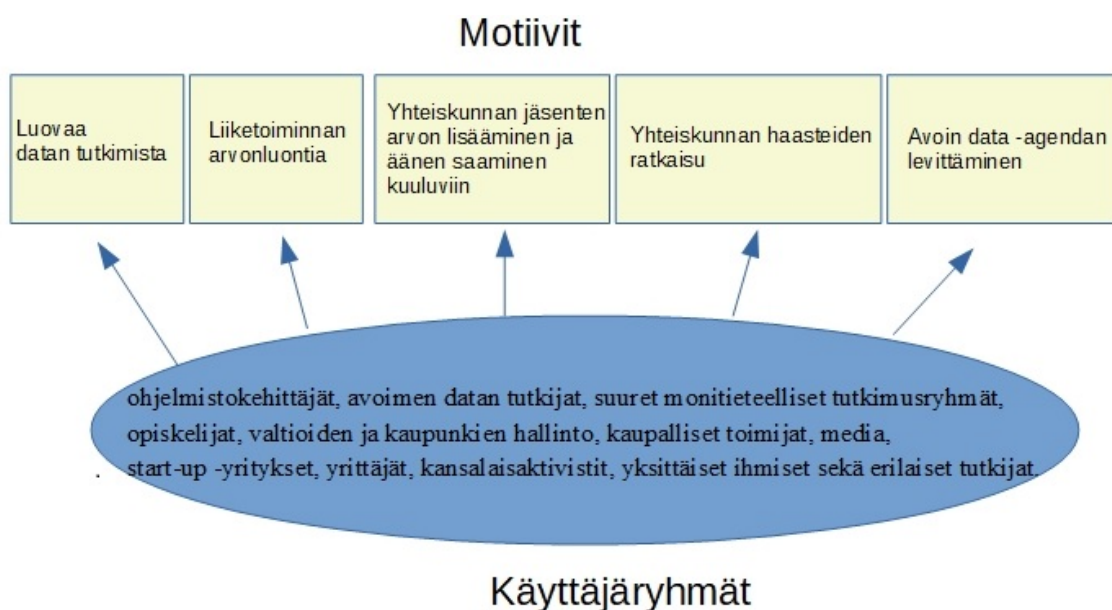
Kuva 8. Ilmastomuutoksen suuraineistojen käyttämisen systeemiarkkitehtuuri (mukaillen Yang et al. 2019).

Erilaisia paikkatietoja on käsitelty Geographical Review -nimisessä lehdessä. Havaintoja ja matkapuhelinten paikkatietodataa (puheluihin liittyviä paikkatietoja) on käytetty tutkittaessa ihmisten liikkeitä Bangladeshissa hirmumyrsky Mahasen aikana. Tutkimuksessa pyrittiin havainnoimaan ihmisten liikkeitä hirmumyrskyn lähestyessä ja rannikkoalueiden evakuointisuunnitelmia tehdessä. Suuraineistanalyysin apuna on käytetty myös tutkijoiden tekemiä havaintoja paikallisten (tyypillisesti kalastusalueiden)

liikkeistä. Tutkimuksessa pyrittiin myös ymmärtämään erilaisten ihmisryhmiä käyttäytymistä ja liikkumista luonnonkatastrofia ennen ja sen aikana. Tutkimuksessa tutkittiin eri ihmisryhmien siirtymistapoja ja siirtymispaikkoja kun hirmumyrsky lähestyi. Vastaavia tutkimuksia on tehty aiemminkin liittyen ihmisten liikkumiseen erilaisista luonnonkatastrofeista johtuen. Tyypillisesti vastaavanlaisissa tutkimuksissa on keskitytty pelkästään suuraineistojen analysointiin eivätkä tutkijat ole tyypillisesti itse menneet paikanpäälle tarkkailemaan ja tekemään havaintoja. Hirmumyrsky Mahasen aiheuttamia ihmisten liikehdintöjä tutkittiin 2.5 kuukautta vuonna 2017. Tutkijat keräsivät havaintoja pääosin haastatteleamalla paikallisia jotta tutkijoille syntyisi parempi ymmärrys tilanteesta. Suuraineistoa käytettiin apuna suunniteltaessa sitä, keitä ihmisiä haastateltiin. Suuraineiston perusteella ihmiset mm. siirtyivät kotitalojensa sekä veneidensä luo. Suuraineistoista ei kuitenkaan voitu selvittää, miksi näin tapahtui ja haastatteleamalla kävi ilmi, että paikalliset ihmiset halusivat olla suojaamassa omaisuuttaan. Lisäksi erityisesti miespuoliset kalastajat liikkuivat hirmumyrskyn aikaan satamassa tarkastamassa laivojensa tilannetta (Boas et al. 2019). Toisin kuin vaikka sosiaalisen median päivitykset, paikkatieto on todennäköisesti hyvinkin järkevä analysointikohde johon ei liity mitään Stephens-Davidowitzin kirjassakin mainittuja vääristymiä (esim. sosiaaliseen mediaan kirjoitetut päivitykset voivat olla vääristyneitä koska ihmiset pyrkivät ehkä alitajuisesti kertomaan vain itselleen suotuisia asioita). Ihmisten ja erilaisten ihmisryhmien liikkeistä on myös hyvin vaikea saada kootusti tietoa muilla tavoin kuin esim. matkapuhelinten paikkatietoa käyttämällä. Tutkimuksessa ei ole otettu kantaa siihen, miten väestö, joka ei omista matkapuhelinta (oletettavasti köyhin väestönosa) liikkui. Suomessa Pelastuslaitoksilla on käytössä Erica-tietojärjestelmä, jossa yhdistellään eri lähteistä saatuja tietoja ja välitetään niitä pelastusyksiköille (Uudenkaupungin Sanomat (2019)). Yksi mielenkiintoinen kehityssuunta voisi olla erilaisten suuraineistojen reaaliaikainen analysointi (esim. matkapuhelinten sijaintitiedot Bangladeshin evakointitutkimuksen tapaan) ja tämän tiedon yhdistäminen esim. pelastusviranomaisten järjestelmiin. Vuonna 2013 Yhdysvaltojen viranomaiset ovat julkaisseet palvelun, joka yhdistää energiasektorin tietoja (voimaloita, siirtolinjoja yms.) sekä tietoja uhkaavista sääilmiöistä reaaliajassa (EIA 2019). Todennäköisesti myös sosiaalisen median datan analysoinnissa pitäisi suuraineistoanalyysien lisäksi huomioida vielä erilaiset kulttuuriset seikat, koska pelkkä data-analyysi ei välttämättä kaikkea paljasta. Netistä löytyy Tobacco Watcher -niminen reaaliaikainen uutisia ja sosiaalista mediaa seuraava palvelu. Tobacco Watcher -sivuston avulla voi seurata esim. tupakkaan liittyvää uutisointia (kieltoja, mainontaa yms.). Sivusto luokittelee erilaisia tupakkaan liittyviä uutisia tulostaa reaaliajassa näistä visualisointeja (TobaccoWatcher 2019). Sivustolla voi myös lukea tupakkaan liittyviä uutisia ja nähdä näiden uutisten luokitteluja. Valitettavasti osa uutisista ei liity tupakkaan millään lailla ja analyysihin kannattaa suhtautua varauksella. Todennäköisesti sosiaalisen median ja uutisten automaattinen luokittelu vaatii vielä työtä. On myös huomattava, että erilaisia suuraineistoja analysoitaessa olisi myös ymmärrettävä dataan liittyvät muut asiat (esim. kulttuurin vaikutus uutisissa ja sosiaalisessa mediassa ja esim. maatalousdataan liittyvät sosioekonomiset ja kulttuurilliset näkökulmat). Geographical Review -julkaisun artikkelissa käytettiin pelkän suuraineistoanalyysin piiloon jäävästä (joka haastattelujen kautta paljastui) tiedosta termiä ”data shadow”.

Government Information Quarterly -julkaisussa on pyritty luokittelemaan erilaisten

suuraineistojen käyttäjäryhmiä sekä selvittämään erilaisten käyttäjäryhmien motiiveja suuraineistojen käyttämiselle ilmastonmuutokseen liittyen. Tutkimuksessa on haastateltu erilaisia julkishallinnon avoimen datan käyttäjiä ja pyritty luokittelemaan käyttäjiä sekä motiiveja. Tutkimuksen mukaan erilaisia suuraineistoja käytetään erilaisiin luoviin kokeiluihin, yritysten arvonaluontiin, yksittäisten käyttäjien palveluihin luomiseen, sosiaalisten haasteiden ratkaisuun sekä avointen aineistojen käyttämisen markkinointiin ja tunnetuksi tekemiseen. Tyypillisiä käyttäjäryhmiä ovat esim. ohjelmistokehittäjät, avoimen datan tutkijat, suuret monitieteelliset tutkimusryhmät, opiskelijat, valtioiden ja kaupunkien hallinto, kaupalliset toimijat, media, start-up -yritykset, yrittäjät, kansalaisaktivistit, yksittäiset ihmiset sekä erilaiset tutkijat. Kun erilaisia suuraineistoja on yhä enemmän saatavilla, suuraineistojen käyttäminen lisää tyypillisesti yhteiskunnan avoimuutta ja hallinnon läpinäkyvyyttä. Toisaalta suuraineistojen käyttäminen vaatii myös IT-alan taitoja joita ei kaikilla kansalaisilla ole (Lassinantti et al. 2019). Erilaista ilmastonmuutokseen liittyvää suuraineistojen analysointi voi tehdä hyvin erilaiset toimijat kuten valtiot, liikeyritykset, yksityiset ihmiset sekä erilaiset harrastelijat. Vapaasti saatavilla olevat erilaiset suuraineistot mahdollistavat myös kattavien havaintojen käyttämisen yksittäisille toimijoille. Tiede lehti uutisoi vuonna 2010 ilmastonmuutoksen kieltäjistä ja epäilijöistä (Tiede 2010), jotka pystyvät myös hyödyntämään avoimien lähteiden suuraineistoja ilmastonmuutoksen liittyvien asioiden epäilemisessä. Epäily ei välttämättä ole huono asia, sillä epäily voi lisätä ihmisten kiinnostusta suuraineistojen analysointiin sekä myös tuoda uusia näkökulmia keskusteluun.



Kuva 9. Avoimen lähteiden suuraineistojen käyttäjä ja motiiveja (Lassinantti et al. 2019).

Journal of Cleaner Production -julkaisussa on käsitelty suuraineistojen käyttämistä kestäväen kehityksen periaatteilla toimivaan tuotantoon sekä tuotantoketjujen hiilipäästöjen vähentämiseen. Tutkimuksessa suositellaan erilaisten käytössä olevien (yritysten ja avointen lähteiden) suuraineistojen yhdistämistä ja hyödyntämistä ”vihreämpien” tuotantoketjujen, henkilöstöhallinnon sekä yleisesti yritysten erilaisten suuraineistojen hallinnan parissa. Suuraineistojen analysointi tällä tavoin voi

tutkimuksen mukaan lisätä yrityksen tehokkuutta. Tehokkuuden lisäksi näiden asioiden hyödyntäminen yrityksen strategiassa lisää yrityksen mahdollisuuksia toimia paremmin kestävämmän kehityksen mukaisesti. Tutkimus nostaa esiin erilaisten reaaliaikaisten suuraineistoanalyysien hyödyntämisen sekä yrityksen eri osien tiedonvälityksen parantamisen. Tutkimus toteutettiin tutustumalla eri yritysten organisaatioihin sekä haastatteleamalla eri yritysten henkilöitä. Tutkimuksessa luotiin matemaattinen malli yrityksen toiminnalle, jonka avulla yrityksen toiminnan tehokkuutta ja kestävä kehityksen mukaista toimintaa arvioitiin numeerisesti (Singh et al. 2019). Voi kuitenkin olla, että yrityksiä ohjaavat enempi lyhyen aikavälin taloudelliset tavoitteet kuin esim. pitkän aikavälin ilmastonsuojelulliset tavoitteet. Ylen uutisen mukaan kuluttajat vaativat kyllä yhä useammin yrityksiltä suurempaa eettistä ja ympäristönsuojelullista vastuuta, mutta tätä vastuuta voi olla hankala luotettavasti todentaa ja osa yrityksistä saattaa pyrkiä tietoisesti harhauttamaan kuluttajia (Yle 2019a). Lisäksi yritysten ympäristöasioista huomioiminen voi riippua hyvin paljon myös toimialasta. Elektroniikkalaitteiden suunnittelua tekevällä yrityksellä voi olla erilaiset intressit ja tarpeet big earth data -tyylisten suuraineistojen käytölle kuin vaikka ydinvoimalaa rakentavalla yritykselle. Ylen uutisen mukaan maailman energiantuotannosta on tehty laskelmia, joiden mukaan siirtyminen uusiutuvasta energianlähteistä uusiutuviin, on tulevaisuudessa kannattavaa (Yle 2019b). Voi olla todennäköisempää, että erilaisia big earth data -aineistoja käyttävätkin tulevaisuudessa enempi suuremmat toimijat (esim. energiateollisuus, valtiot yms.) kuin pienet yritykset. Pienten yritysten toiminta voi olla lähinnä paikallista ja lyhyen aikavälin toimintaan keskittynyttä. Samalla tavalla aiemmin tässä lopputyössä oli Google Trends -palvelun hakutietoja ilmastomuutokseen liittyvistä hauista. Ihmiset olivat hakeneet ”climate change” -hakusanoilla Googlesta tietoa lähinnä silloin, kun haku palveli ihmisten omia etuja (tietoa haettiin ilmastomuutoksen aiheuttamien helteiden aikana ja tietoa ei haettu jouluna) (Google Trends 2019). Yrityksissä on kuitenkin töissä ihan samoja ihmisiä, jolloin yrityksen oma (lyhytaikainen) taloudellinen etu asetetaan ehkä tärkeämmäksi kuin ilmastomuutoksen torjuminen. Ylen teettämän kyselyn mukaan joka toinen kyselyyn vastannut yritys on reagoinut jollain tasolla ilmastomuutoksen aiheuttamiin mahdollisiin tulevaisuuden uhkiin (Yle 2019c). Journal of Cleaner Production -julkaisun tutkimus on esimerkki hyvin korkean tason mallinnuksesta ja tutkimuksen loppupäätelmien (suuraineistojen analyysien hyödyntäminen yrityksen toiminnassa) toteuttaminen käytännössä voi olla hyvin hankala toteuttaa. Todennäköisesti toteutus on lisäksi hieman erilainen joka yrityksessä sekä eri toimialoilla. Yksi mahdollinen käyttökohde voisi tämäntyyppisille analyyseille olla esim. kuljetusten reittien optimointi eli ns. vihreä logistiikka, jota on käsitelty mm. Haaga-Helian Ammattikorkeakoulussa tehdyssä opinnäytetyössä (Kinnunen 2016). Ongelmana voi olla se, että suuraineistojen analysoinnista saadut tulokset eivät välttämättä ole aina liiketoiminnan kannalta hyödynnettävissä (esim. logistiikassa kaikki tavara on kuljetettava ja vaikka ajoreittejä optimoitaisiin, niin ilmaston kannalta suurempi hyöty voi tulla esim. uusiutuviin polttoaineisiin siirtymisistä kuin monimutkaisesta ajoreittilaskennasta).

Global Sustainability -lehdessä on käsitelty suurkaupunkien (Manhattan, Berliini, New Delhi) ja ilmastomuutoksen välisiä yhteyksiä. Tutkimuksessa on käsitelty ilmastodatan (esim. lämpötilat, satelliittikuvista saatu tieto yms.) ja sosiaalisen median datan yhdistämistä, sekä tästä syntyneen datan analysointia. Jotta suuraineiston käsittely ja analysointi olisi tehokkaampaa, datan olisi syytä olla skaalautuvaa (samaa dataa voidaan

hyödyntää useiden kaupunkien tilanteen analysoinnissa), koneoppimista olisi syytä hyödyntää (analysointi on tehokkaampaa ja ihmisten yksilöllisyyteen liittyvät asiat tulee huomioida asianmukaisesti) sekä analyysien tuloksista pitäisi saada tietoa ilmastonmuutokseen liittyen (esim. käyttämällä yhdenmukaisia analyysejä). Jotta toiminta olisi tehokkaampaa, datan yhtenäistäminen (rajapinnat, eri kaupunkien ratkaisujen soveltuminen muihin kaupunkeihin yms.) ja standardisoitu käyttö on järkevää. Kun yhden kaupungin hiilidioksidipäästöjen lähteet on selvitetty ja ratkaisut ilmastoystävällisempään kaupunkisuunnitteluun ovat selvillä, samat ratkaisut pitäisi olla siirrettävissä muihinkin kaupunkeihin. Samalla kun eri kaupunkien hiilidioksidipäästöjä ja muuta kaupungeista kerättyjä suuraineistoja, voidaan saada tietoa ilmastonmuutoksen vaikutuksista ja etenemisestä. Tutkimus käyttää termiä ”upscaling” kuvaamaan sitä, miten yhdessä paikassa (yksi kaupunki) tehtyjä ilmastonmuutokseen liittyviä toimenpiteitä kopioidaan käytettäväksi laajemmin (esim. toisessa kaupungissa) (Creutzig et al. 2019). Ehkä tämentyylisistä (skaalautuvista) analyyseistä voitaisiin päätellä, mitkä asiat ovat erityisen ongelmallisia ilmastonmuutoksen suhteen (esim. päästöt, infrastruktuurin heikkoudet esim. äärisäille yms.), miten eri kaupungeissa on kyseiset asiat hoidettu (ts. kopioidaan toimivat ratkaisut muuallekin). Kaupungistuminen on yksi megatrendeistä, ja mikäli kaupungistuminen jatkuu tulevaisuudessakin, yhä suurempi osa ihmisistä tulee todennäköisesti asumaan kaupungeissa. Tästä voi seurata se, että kaupunkien merkitys ilmastonmuutoksessa kasvaa ja kaupunkien tuottama tietomäärä lisääntyy. Tutkimuksessa oli käsitelty hyvin monimutkaisiakin teknologisia ratkaisuja (esim. liikenteen kelikameroiden kuvien analysointi automaattisesti), mutta on myös muistettava, että kaikki suuraineistot eivät välttämättä vaadi monimutkaista teknologiaa. Suomessa esim. Luonto-Liitto järjestää keväisin ns. kevätseurannan, jossa kevään merkkejä havainnoidaan (päivämäärät ja havainnointipaikat kirjataan ylös, milloin mikäkin kevään merkki on havaittu ja tietoja kevään etenemisestä on kerätty usealta vuodelta) (Luonto-Liitto 2019). Jos kaupunkien tuottamat datamassat paisuvat yhä edelleen, niin ehkä yksi kehityssuunta voisi olla Luonto-Liiton kevätseurannassa käytettävien yksinkertaisten ilmastonmuutosindikaattorien havainnointi. Tällöin dataa kerättäisiin vähemmän mutta data olisi ehkä jalostuneempaa. Vaikka datan määrää kasvatetaan (nykyään erilaiset kodinkoneetkin saattavat alkaa tuottaa sitä dataa), niin voi olla kyseenalaista, kuinka paljon tietoa datasta saadaan analysoitua (kasvava datamäärä ei välttämättä tarkoita kasvavaa informaation määrää). Global Sustainability -lehden tutkimuksen ratkaisujen (skaalautuvat suuaineistoanalyysit eri kaupunkien suunnittelussa) toteuttaminen voi olla kuitenkin haasteellista erilaisten kulttuuristen ja muiden kansallisten piirteiden (esim. autoilun hinta voi olla erilainen eri maissa ja suhtautuminen yksityisautoiluun voi vaihdella eri kaupungeissa) vuoksi. Lisäksi eri maissa ihmisten suhtautuminen ilmastonmuutokseen hillitsemiseen ja talouskasvuun voi vaihdella, jolloin toisesta kaupungista kopioidut ratkaisut eivät välttämättä toimi toisessa kaupungissa sellaisenaan. Pasi Toiviaisen kirjoittamassa jutussa on esim. arveltu, että tiheästi rakennetussa kaupungissa ihmisten yksityisautoilun päästöt ovat huomattavasti pienemmät kuin laajalle alueelle rakennetussa kaupungissa (Yle 2015a). Jos esim. edellä mainittua yksityisautoilun ongelmaa ratkaistaan kopioimalla joukkoliikennemalli toisesta kaupungista, niin ihmiset eivät välttämättä halua käyttää joukkoliikennettä vaikka erilaiset (suuraineisto)analyysit joukkoliikenteen käyttämistä suosittelisivatkin. Jyväskylän Yliopistossa vuonna 2018 tehdyssä, kaupunkisuunnittelun digitalisaatiota käsittelevässä kandidaatintutkielmassa on yhtenä näkökulmana suuraineistojen

käyttäminen kaupunkisuunnittelussa (Kronholm 2018). Lienee myös todennäköistä, että erilaisen datan määrän lisääntyminen mahdollistaa tulevaisuudessakin suuraineistojen hyödyntämisen kaupunkisuunnittelussa.

Cities -lehdessä julkaistussa artikkelissa käsitellään erityisesti kaupunkiympäristöstä kerättyjä suuraineistoja. Tämän tiedon hyödyntämisestä käytetään nimitystä ”smart cities”. Erilaisten sensorien, IoT-laitteiden avulla pystytään keräämään suuri määrä erilaista tietoa. Tutkimuksen mukaan tyypillisesti erilaisissa tämentyylisiä suuraineistoja analyysoivissa ja käyttävissä ratkaisuissa painotetaan enempi erilaisten uusien teknologioiden ja apuvälineiden käyttämiseen. Insinöörimäisen lähestymistavan lisäksi pitäisi huomioida myös erilaiset kulttuurilliset ja hallinnolliset seikat. Tyypillisesti erilaisia tekoälysovelluksia käytetään prosessoimaan erilaisia suuraineistoja. Tyypilliset suuraineistojen sovellukset liittyvät koulutukseen, ympäristönsuojeluun, energiaan, terveydenhuoltoon, poliittisiin päätöksiin (policy), taloudellisiin asioihin, IT-järjestelmiin liittyviin toimintoihin, liikkuvuuteen, kestävään kehitykseen, IoT-sovelluksiin tai suuraineistojen liittyvään laskentaan ja analysointiin yleensäkin (Allama et al. 2019). Kun erilaisilla sensoreilla kerätään suuraineistoja, voidaan saatua tietoa käyttää hyvin monella tavalla ja myös tähän liittyvät eettiset näkökulmat olisi hyvä huomioida. Ylen uutisen mukaan Kiina käyttää erilaista kasvojentunnistuksen ja muiden verkottuneiden laitteistojen muodostamaa tietomassa pisteyttääkseen kansalaisiaan (Yle 2018a). Erilaisten ilmastomuutokseen liittyvien suuraineistojen analyysien tulokset voivat tarkoittaa myös erilaisten ihmisryhmien ja yritysten suosimista tai etujen leikkaamista. Ylen uutisten mukaan suomalainen Neste on kehittänyt edistyksellistä biopolttoaineteknologiaa (Yle 2019d) ja esim. kaupunkien ja valtioiden päätökset siirtyä fossiilisista energiamuodoista biopolttoaineisiin hyödyntäisivät Nesteen kaltaisia toimijoita ja vastaavasti heikentäisivät perinteisen öljyteollisuuden toimijoita. Myös erilaisten ilmastomuutokseen liittyvien suuraineistojen analysoinnin lopputulokset (erilaiset tulevaisuuden uhkakuvat) voivat aiheuttaa ihmisissä psykologista kuormitusta ja stressiä. Ylen uutisen mukaan osa suomalaisista nuorista kokee ilmastomuutoksen ja siihen liittyvät asiat ahdistavina ja tästä käytetään termiä ”ilmastoahdistus” (Yle 2019e). Lisäksi avoimen datan aineistoja voidaan periaatteessa vääristellä tai jättää olennaisia tietoja pois. Ylen uutisen mukaan yksittäiset ihmiset eivät välttämättä edes halua tietää omiin elämäntapoihin liittyvää ilmastokuormitusta (Yle 2019f). Tällöin, vaikka käytettävissä olisi kattavat suuraineistot ja analyysit, yksittäiset ihmiset eivät välttämättä halua reagoida tutkimustuloksiin. Elinkeinoelämän Valtuuskunnan kesäkuussa 2019 julkaistun haastattelututkimuksen mukaan suomalaisten halukkuus vähentää kulutusta ilmastomuutoksesta johtuen on vähentynyt viime vuosina (Elinkeinoelämän Valtuuskunta 2019). Vaikka erilaisia ilmastoon liittyviä suuraineistoja todennäköisesti nykyäänkin käytetään kaupunkisuunnittelussa apuna, teknisen data-analyysin lisäksi ilmastomuutos vaatii myös ihmisiltä elämäntapojen muuttamista.

Erilaisten käytettävissä olevien julkisten lähteiden suuraineistojen määrä on nykyään kasvanut huomattavasti. Erilaisia aineistoja yhdistettäessä ongelmia voi tulla liittyen data-aineistojen esitystapoihin, kerättyjen data-aineistojen laajuuteen (esim. data-aineisto voi olla kerättyä vain tietyltä maantieteelliseltä alueelta) sekä käytettävään terminologiaan (data voi kuvata erilaisia asioita ja eri aineistojen yhdistämisessä tämä on huomioitavat). BLoS Biology -lehdessä julkaistussa artikkelissa on käsitelty erityisesti biodiversiteetteihin liittyviä data-aineistoja. Tyypillisesti biodiversiteetti on

käsitteenä hyvin laaja. Tästä syystä biodiversiteetteihin liittyvissä data-aineistoissa pitäisi määritellä, mitä biodiversiteetin osa-aluetta kerätty data koskee ja mikä on datan resoluutio (esim. kuinka usein aineiston havainnot on tehty). Biodiversiteetteihin liittyntä dataa kerätessä aineisto on myös saatettu koostaa jonkun ongelman perusteella (esim. on kerätty näytteitä tietyistä paikoista ja tietyistä kasveista, jotta jotain eksaktia tutkimusongelmaa on voitu lähteä ratkaisemaan). Tyypillistä onkin, että biodiversiteettidataa käsiteltäessä joudutaan yhdistelemään erilaisia suuraineistoja. Lisäksi suuraineistosta saattaa puuttua dataa (esim. dataa ei ole joka vuonna kerätty) jolloin puuttuvat datat täytyy jollain tavalla (esim. interpoloimalla) lisätä suuraineistoon (König et al. 2019). Tyypillinen tämän aihealueen ongelma voisi olla vaikka tietyistä kalalajista kerätyt havainnot tietyllä maantieteellisellä alueella. Jossain toisessa tutkimuksessa voitaisiin tutkia toista kalalajia, ja ehkä tällöin olisi tarpeen yhdistää aiemmin tehty aineisto tähän tutkimukseen. Tällöin ongelmaksi voi tulla se, miten kalalajien esiintymistä on aiemmin kuvattu ja miten aiemman tutkimuksen mahdollisesti ajallisesti eri aikoina kerättyä dataa voidaan hyödyntää uudessa tutkimuksessa. Yle käsitteli uutisessaan ilmastonmuutoksen aiheuttamaa maa- ja merieläinten sukupuuttouhkaa ja tutkimuksen mukaan merieläimillä on kaksinkertainen sukupuuton riski (Yle 2019g). Todennäköisesti on olemassa paljon erilaisista eläimistä kerättyjä aineistoja, joita on tarpeen analysoida ja yhdistellä vastaaviin, toisista eläimistä kerättyihin tietoihin. Lopputuloksena voi helposti olla erilaisia tiedonhallintaan liittyviä haasteita kun erilaisia ja erityyppisiä aineistoja yhdistellään. Todennäköisesti tarvitaan vuoropuhelua ei alojen tutkijoiden kesken, jotta erilaisista aineistoista saadaan paras lopputulos.

Nature-lehdessä julkaistussa artikkelissa käsiteltiin koneoppimisen hyödyntämistä erilaisissa maapallon ilmastoon ja muuhun vastaavaan mittaustietoon liittyvissä tutkimuksissa. Tutkimuksessa pohdittiin sitä, miten erilaiset kausivaihtelut voidaan erotella kerätyistä datasta sekä huomioida erilaisissa pitkän aikavälin ennustuksissa. Artikkelin käsitteli teknisiä ongelmia koneoppimiseen ja tämän tyylisten suuraineistojen analysointiin liittyen, sekä ehdotti erilaisia parannuksia tähän liittyvään tekniseen laskentaan. Lisäksi tarkasteltiin perinteisiä koneoppimiseen liittyviä tehtäviä (esim. kuvien sisällön tunnistaminen) ja näitä tehtäviä verrattiin vastaaviin luonnosta mittauksilla kerättäviin suuraineistojen analyysitehtäviin (esim. sääsatelliittikuvan piirteiden etsintä). Artikkelin mukaan maapallosta ja luonnosta dataa kerätessä data-analyysien suurimpia ongelmia ovat:

- datan tulkittavuus ja ymmärrettävyys (esim. datan visualisointi voi auttaa)
- datan fyysinen yhtenäisyys (consistency, kuinka kattavasti havainnot kuvaavat mitattavaa asiaa)
- datan monimutkaisuus ja epävarmuus
- datan nimeämisen rajoitteet (ts. mitattavia suureita ja tunnuslukuja voi olla paljon)
- data-analyysien vaatima laskennallinen teho

Parannusehdotuksina esitetään erilaisten simulaatiomallien käyttämistä, joissa on tietty määrä ymmärrettäviä parametreja jotka huomioivat myös erilaisia epävarmuuksia (uncertainty), jotka liittyvät esim. erilaisten luonnonilmiöiden monimutkaisuuteen (Reichstein et al. 2019). Artikkelin mukaan suuraineistojen analysointi ja erilaiset mallintamiset vaativat hyvin syvällistä tietoa koneoppimisesta ja koneoppimiseen liittyvistä tekniikoista. Näiden tietojen lisäksi tarvitaan todennäköisesti myös tietoa itse

tutkittavista asioista (esim. ilmaston käyttäytymisen tulkinta). On melko selvää, että esim. ilmastonmuutoksen tutkimisessa tarvitaan hyvin monenlaisia ihmisiä, joilla on osaamista eri aloilta. Artikkelit tuntuvat myös viestivän siitä, että vaikka dataa on hyvin paljon ja käytettävät menetelmät ovat hyvin teknisiä (koneoppimista ja signaalinkäsittelyn algoritmeja), niin oletettavasti data-analyysien pohjalta saatujen mallien pitäisi olla mahdollisimman ymmärrettäviä ja havainnollisia. Yhtenä tapana lisätä erilaisten mallien ymmärrettävyyttä voisi olla erilaisten visualisointien käyttäminen, jossa käyttäjä voi itse säätää erilaisia parametreja ja havainnoida lopputulosta.

Applied Energy -julkaisussa on käsitelty erityisesti suuraineistojen käyttöä erilaisissa energiateollisuuden hankkeissa. Energiateollisuuteen liittyvissä suuraineistoanalyysissä saatetaan joutua usein käyttämään satunnaista kerättyä dataa, mitä sattuu olemaan saatavilla (erilaisten mallinnusten tarvitsemaa dataa ei välttämättä ole saatavilla). Tällöin erilaisia malleja ja analyysiejä varten saatetaan joutua luomaan puuttuvaa dataa. Tutkimuksen mukaan suuraineistojen tehokkaalla käyttämisellä ja erilaisten mallien luomisella, voidaan tunnistaa erilaisia energiantuotannon muutoin piiloon jääviä uhkia ja asioita. Erilaisilla malleilla voidaan mallintaa energiateollisuuden loppukäyttäjien tarpeita, energiantuotantojärjestelmän tehokkuutta ja käytettävien energialähteiden valintaa. Tyypillistä on, että tämäntyyllisiä asioita suunniteltaessa ei ole käytettävissä tarpeeksi dataa, jolloin erilaiset suunnittelussa tarvittavat tiedot jäävät saamatta. Tutkimuksessa arvellaan, että käyttämällä erilaisia suuraineistoja, voidaan myös tämäntyylliset ongelmat ratkaista. Tutkimuksessa käydään läpi myös suuraineistojen tyypillisiä käyttökohteita, joita ovat nykyään esim. terveydenhuolto, kuluttajatrendien analysointi sekä muut, erityisesti kaupallisiin tarkoituksiin tehdyt suuraineistoanalyysit (Li et al. 2019). Hyödyntämällä ja yhdistelemällä erilaisia datalähteitä, sekä analysoimalla näin saatua dataa, voidaan monimutkaisia järjestelmiä ymmärtää paremmin. Siemens on kehittänyt ns. virtuaalivoimalaa, jossa sähköverkosta tulevaa sähköverkon kuormitusdataa voidaan käyttää sähköön loppukäyttäjällä, jonka kulutusta voidaan kulutushuippujen kohdalla vähentää (Siemens 2019). Siemensin virtuaalivoimalassa yhdistyvät kuluttajan energiantarpeen sääntely verkon tuotantokapasiteettiin ja lopputuloksena energiantuottajan ei tarvitse ottaa käyttöön kalliimpaa varavoimaa kulutushuippujen tasaamiseksi. Samalla todennäköisesti saadaan kerättyä reaaliaikaista tietoa sille, miten eri tilanteissa asiakkaat käyttävät energiaa ja energiantuotantoa suunnitellessa voidaan tätä tietoa käyttää hyödyksi. Tekesin vuonna 2017 julkaistussa taustaraportissa on visioitu tulevaisuuden energiajärjestelmä vuonna 2035, jossa yhtenä osana on datakeskuksia sekä erilaista tiedonsiirtoa liittyen järjestelmän ohjaamiseen ja sähkömarkkinoihin (Tekes 2017, sivu 14). Lienee hyvin todennäköisestä, että myös energia-alalla erilaisten suuraineistojen kerääminen ja luominen on kasvava trendi. Ilmastonmuutos saattaa aiheuttaa tarpeita tehdä energiatehokkaampia energiajärjestelmiä jotka pystyvät reagoimaan tehokkaammin ympäristön, luonnon ja markkinoiden muutoksiin.

Geography and Environment -lehdessä julkaisussa on käsitelty ilmastonmuutokseen liittyvien data-analyysien visualisointia. Tutkimuksen mukaan usein ongelmana on se, miten erilaisten mallinnusten ja etenkin visualisointien lopputulokset muutetaan paikallisten yhteisöjen käytännön toiminnaksi (ts. mitä paikallisten yhteisöjen pitäisi tehdä käytännössä jotta ilmastonmuutokseen voitaisiin reagoida). Mikäli globaalisti

kerätyistä suuraineistosta tehdään tiettyä maantieteellistä aluetta koskevia analyyskejä, globaali aineisto ei välttämättä huomioi tarpeeksi kyseisen alueen (datan) erikoispiirteitä ja alueen ilmiöt saattavat hukkaa suurempien globaalien ilmiöiden alle. Yhtenä ratkaisuna on käyttää paikallisissa visualisoinneissa esim. paikallisten ihmisten tekemiä havaintoja. Tutkimuksessa käytetään termiä ”downscaling” globaalien ilmastomuutosmallinnusten soveltamisesta paikalliseen yhteisöön/alueeseen. Vastaavasti termi ”upscaling” viittaa päinvastaiseen tilanteeseen, jossa paikallisia ilmastomuutokseen liittyviä ratkaisukeinoja sovelletaan globaalisti tai maantieteellisesti eri paikassa (Schneider et al. 2019). Vaikka tutkimus on hyvin teoreettinen (eikä matemaattinen), on tutkimuksessa kuvailtu haaste (miten globaalit datan visualisoinnit muunnetaan siten, että paikalliset ihmiset voivat käyttää noita paikallisessa toiminnassa) todennäköisesti hyvinkin konkreettinen. Erilaisia ilmastomuutosta koskevia analyyskejä ja visualisointeja on esim. uutisissa hyvin paljon, mutta yksittäisen ihmisen tai paikallisen yhteisön voi olla hankala hyödyntää niitä.

Käsitellyt tutkimukset toivat esiin hyvin samansuuntaisia ongelmia data-aineistojen käsittelyssä, kuin mitä Kristiina Heinosen lopputyössä oli käsitelty (Heinonen 2016). Data-aineistojen ongelmat ovat hyvin samanlaisia tutkimuskohteesta (esim. ilmastomuutos) riippumatta.

Ehkä tutkimusten perusteella yhtenä merkittävimpänä erona perinteiseen (yritysten käyttämään, pääosin lyhytaikaiseen kaupalliseen tarkoitukseen) data-analytiikkaan oli ilmastoon liittyvien suuraineistojen tulkinnan vaikeus ja erilaisten ilmiöiden vuorovaikutussuhteiden ymmärtäminen. Ilmastomuutoksen tutkimiseen liittyvät suuraineistot ovat usein suurten julkisten toimijoiden (kuten valtiot) keräämiä kun taas perinteinen yritysten data-analytiikka saattaa hyödyntää pääosin oman yrityksen keräämää dataa. Vanhemmat ilmastomuutoksen suuraineistot saattavat olla myös ei-digitaalisia ja tarkat tiedot siitä, miten data on kerätty saattavat puuttua.

Smart Data Collective -verkkosivuilla julkaistun artikkelin mukaan suuraineistoja analysoimalla voidaan paremmin määritellä eri asioiden hiilijalanjälki, hyödyntää tehokkaammin olemassa olevia resursseja, ennakoida mahdollisia tulevaisuuden äärisäitä sekä päättää, mitkä asiat vaativat välittömintä huomiota (SmartDataCollective 2019). Tämän tyylliset analyysit ovat hyvin haastavia ja jossain määrin epämääräisiä (välttämättä ei tiedetä, mitä etsitään). Käsitellyissä, vuonna 2019 julkaistuissa ilmastomuutoksen suuraineistoja käsittelevissä tutkimuksissa tunnistettiin erityisesti ilmastomuutokseen liittyvien suuraineistojen haasteita. Näiden haasteiden ratkaisuun tarjottiin tutkimusten mukaan erilaisia teknisiä (matematiikkaan perustuvia) että myös tutkimuskohdespesifisiä (esim. paremman ymmärryksen hankkimista tutkittavaan ilmiöön).

4.2 Riskienhallintaan liittyviä tutkimuksia

Tyypillisesti ilmastomuutos saatetaan nähdä riskienhallinnan näkökulmasta ja erilaisten kohonneiden riskien kautta. Suuraineistoja analysoimalla voidaan saada käsitys siitä, millaiset tekijät vaikuttavat erilaisten haitallisten tapahtumien tapahtumistodennäköisyyksiin. Tämä kappale käsittelee ilmastomuutokseen liittyvien suuraineistojen käyttämistä riskienhallintaan liittyvissä tutkimuksissa.

Annals of The New York Academy of Sciences -julkaisussa julkaistussa tutkimuksessa on tutkittu pitkän aikavälin tilastotietojen ja ilmastodatan käyttämistä ilmastomuutoksen äärisäiden aiheuttamien riskien mallintamisessa. Tutkimuksessa on myös pyritty hahmottelemaan erilaisten äärisäiden aiheuttamia worst-case -skenaarioita liittyen esim. hirmumyrskyjen vaikutuksiin ja ilmaston lämpenemisen maanviljelylle aiheuttamiin vaikutuksiin. Riskien ennustamisessa on tyypillisesti käytetty useita datalähteitä, joiden avulla esim. säätä on mallinnettu. Lisäksi analyysihin on yhdistetty tietoja riskien aiheuttamista vaikutuksista. Erilaista säädataa on ollut saatavilla tyypillisesti 1950-luvulta lähtien ja joissain maissa säätietoja on kerätty jo 1900-luvun alussa. Tätä vanhempaakin säädataa on olemassa ja säädataa on kerätty liittyen esim. maanviljelykseen ja myrskyjen esiintymisiin. Tyypillinen (tyypillinen tutkimuksissa käytettävä aineisto) ilmastodataa sisältävä suuraineisto kattaa dataa 50-70 vuoden ajalta (Brönnimann et al. 2019). Tämän tyyliä riskianalyysille on todennäköisesti tarvetta kun valtiot varautuvat ilmastomuutoksen vaikutuksiin. Kaupallisista toimijoista todennäköisesti erityisesti vakuutusyhtiöt tarvitsevat erilaisiin riskeihin liittyviä simulaatietietoja. Kuitenkin Ylen uutisen mukaan Suomessa lähinnä valtion yksittäiset virastot ovat jollain tasolla varautuneet ilmastomuutokseen ja yksityisistä yrityksistä harva on varautunut ilmastomuutokseen (suomalaisista pörssiyrityksistä noin seitsemän prosenttia on huomionnut ilmastomuutoksen omassa strategiassaan, erityisesti Neste) (Yle 2019h). Ehkä yritysten strategioita ei suunnitella kovin pitkälle ja ehkä siksi yritykset olettavat, että ilmastomuutokseen pystytään reagoimaan hitaasti muiden hitaasti etenevien muutosten mukaan. Suomessa valtion virastoista esim. Maanmittauslaitos on ollut toteuttamassa laserkeilauksella kerättäviin tietoihin perustuvaa tulvien ennakointia (Yle 2018b). Ylen uutisen mukaan Lapin metsälampien pohjaan on aikojen kuluessa jäänyt puunrunkoja, joita ajoittamalla ja vuosirenkaita tutkimalla voidaan saada tietoa tuhansien vuosien ajalta (Yle 2015b). Tämän tyyllisen suuraineiston muuttaminen digitaaliseen muotoon voi olla hidasta, mutta samalla voidaan saada pidemmältä väliltä tietoa ilmastosta.

Hieman samantyylinen tutkimus on julkaistu Geoscientific Model Development -julkaisussa. Tutkimuksessa käytettiin hyväksi erilaisia avoimesti saatavilla olevia säähän liittyviä suuraineistoja. Sääaineistoissa oli tietoja erilaisista äärisäistä kuten tulvien ja myrskyjen esiintymisistä ympäri maailmaa vuosina 1979-2005. Tutkimuksessa käytettiin koneoppimisalgoritmeja, joilla käytetty matemaattinen malli opetettiin tunnistamaan sääolosuhteita, joissa äärisäitä voi esiintyä. Äärisäiden havainnointialgoritmi etsi tiettyjä piirteitä suuraineistoista ja muodosti näiden piirteiden perustella todennäköisyyksiä äärimmäisten sääilmiöiden esiintymiselle eri alueilla (Muszynski et al. 2019). Tutkimus on hyvä esimerkki siitä, miten suurten data-aineistojen käyttämistä voidaan automatisoida ja tehdä erilaisia ennusteita. Keräämällä reaaliaikaista dataa, voidaan ennusteita tehdä myös reaaliaikaisesti. Erilaisissa malleissa ongelmia voi tuottaa myös se, että kattavan data-aineiston kerääminen ja analysointi voi kestää pitkän ajan. Lisäksi mikäli ei tiedetä, että mikä data on oleellista tutkittavan asian kannalta, saattaa kerätty data osoittautua lopulta hyödyttömäksi. Maapallon ilmastossa on myös monenlaisia takaisinkytkentöjä. Ylen uutisen mukaan pohjoisten alueiden (arktisen) lämpeneminen vaikuttaa sekä paikalliseen (arktiseen) että myös globaaliin ilmastoon (Yle 2018c). Tämän tyylisten takaisinkytkentöjen mallintaminen voi olla hankalaa ja voi ehkä vaatia vieläkin pidemmältä ajalta tietoa ilmastosta (kuin mitä

tutkimuksessa oli käytetty).

Hieman samantapainen mallinnus on tehty Environmental and Resource Economics -lehdessä julkaistussa artikkelissa, jossa on tutkittu metsäpalojen esiintymistä (laajuus ja esiintymistiheys) Italiassa vuosina 2000-2011. Tutkimuksessa on yhdistetty sekä alueen säätietoja (lämpötilavaihtelut, sademäärä yms.) että myös Italian eri alueiden sosio-ekonomisia tietoja (vuoden tarkkuudella, esim. alueilla olevat maatilojen eläimet ja alueella harjoitettava maanviljely sekä alueen elintaso). Erilaisilla tilastollisilla menetelmillä on mallinnettu kerätyn datan riippuvuuksia liittyen metsäpaloihin. Lisäksi tutkimuksessa oli tehty ennuste Italian eri alueiden metsäpalojen todennäköisyyksille ajalle 2016-2035 (Michetti et al. 2019). Tutkimus on siitä hyvä, että siinä otetaan huomioon ilmaston lisäksi myös sosioekonomiset asiat. Todennäköisesti metsäpaloihin ja muihin riskienhallintaan liittyviin asioihin liittyy erilaisia monimutkaisia riippuvuussuhteita, joiden määrittelemine voi olla hyvin haastavaa. Ylen uutisen mukaan Ruotsissa on viime aikoina syttynyt metsäpaloja kertakäyttögrillien käytöstä johtuen (Yle 2019i). Erilaisia aineistoja analysoimalla tilastotieteen menetelmillä, voidaan hahmottaa erilaisia riskitekijöitä. Ongelma voi kuitenkin olla, että erilaisia aineistoja on hyvin paljon ja kaikkien keskinäinen analysointi voi olla hyvin haastavaa. Lisäksi analysoinnissa lieene syytä huomioida, onko löydetty riippuvuussuhteet todellisia vai onko kyse sattumasta. Italian metsäpalotutkimuksessa aineistoa oli käytössä vuosilta 2000-2011 ja yhdentoista vuoden aineistoon voi vaikuttaa myös satunnaiset tekijät (todennäköisesti pidemmän aikavälin ilmastodata toisi luotettavampia tuloksia). Metsäpalojen riskianalyysillä on myöskin todellista tarvetta liittyen pelastuslaitosten toiminnan suunnitteluun ja maatalouteen.

International Journal of Safety and Security -julkaisussa on julkaistu tutkimus, missä seismiseen toimintaan liittyviä suuraineistoja on tutkittu, analysoitu sekä mallinnettu. Tutkimuksessa on tutkittu erityisesti mannerlaattojen keskiosien maanjäristyksiin liittyviä asioita Kanadan Montrealin alueella. Lopputuloksena on saatu kartta, jossa esitetään erilaisten voimakkaiden maanjäristysten esiintymistodennäköisyydet Montrealin alueella, jota esim. vakuutusyhtiöt ja rakennuttajat voivat hyödyntää. Tutkimuksessa on käytetty kaupallista ohjelmistoa (avoimen lähdekoodin EqHaz1) ja tutkimusryhmän tekemiä omia mallinnuksia (Tiampo et al. 2019). Vaikka käytettävissä on yhä enempi dataa, ihmiskunnan ymmärrys ilmastomuutoksesta voi olla rajoittunutta ja parhaimmillaan mallit eivät välttämättä osaa luoda puuttuvaa tietoa. Riskienhallinnan kannalta pahimpia riskejä ovat todennäköisesti riskit, joita ei osata ennustaa. Ylen Uutisen mukaan esim. vuonna 2018 Indonesian Sulawesin saarelle iskenyt tsunami oli seurausta maanjäristyksestä, jonka ei pitänyt olla alueella mahdollinen (Yle 2018d). Tutkimuksessa käytetty ohjelmisto on vapaasti ladattavista netistä ja lisäksi aineiston mukana tulee suuraineistoja liittyen maanjäristyksiin. Tutkimus vaikutti enemmän teknillisen laskennan demolta ja todennäköisesti tutkimukseen olisi kannattanut yhdistää muita suuraineistoja tai asiantuntijahaastatteluja, jolloin analyysissä mahdollisesti piiloon jäävät näkökohdat olisi huomioitu.

Climate Change -julkaisussa on tutkittu ilmastomuutoksen vaikutusta Himalajan alueen maataloudelle. Tutkimuksessa käytettiin suuaineistona alueen lämpötilatietoja vuosilta 1951-2013 sekä alueen sademäärädataa vuosilta 1901-2013. Lisäksi käytössä

oli tilastotietoja liittyen alueen maanviljelijöihin ja maanviljelijät oli luokiteltu viiteen eri luokkaan. Tutkimuksessa tutkittiin ilmastomuutoksen vaikutusta eri maanviljelijätyyppien satoon ja sadoista saataviin tuloihin. Tutkimuksen mukaan ilmastomuutos vaikutti joihin maanviljelijätyyppeihin voimakkaammin ja saatua tietoa voidaan hyödyntää ilmastomuutoksen varautumisessa esim. lainsäädännöllisin ja erilaisten tukien avulla. Maanviljelijöihin liittyvä data (varallisuus, viljeltävän maan määrä, koulutustaso yms.) oli kerätty haastattelemalla alueen maanviljelijöitä. Tutkimus yhdisti sosioekonomisen näkökulman ja ilmastomuutoksen riskien arvioinnin eri maanviljelijäryhmille (Shukla et al. 2019). Tutkimus on hyvä esimerkki siitä, miten eri aineistoja (suuraineistot ja haastattelut) voidaan yhdistää ja saada kattavampi ymmärrys lopputuloksesta kuin mikä saataisi jos pelkästään suuraineistoja analysoitaisiin. Maanviljelijöistä ei välttämättä ole olemassa valmiina kattavia tietoja, joten haastattelut saattavat olla järkevin keino. Ylen uutisoi sateiden lisääntymisestä Keskilännen pelloilla ja uutisen mukaan syynä on ilmastomuutoksen aiheuttama lämpeneminen mikä lisää veden haihtumista Meksikonlahdella (Yle 2019j). Ylen uutinen perustuu lähinnä alueella tehtyihin havaintoihin ja todennäköisesti lisääntyneet sateet aiheuttavat erilaisia ongelmia erityyppisille viljelijöille.

Vuonna 2019 julkaistussa tutkimuksessa on käytetty Wall Street Journal -lehdessä esiintyviä ilmastomuutokseen liittyviä avainsanoja erilaisten uutisten ja valtion tiedotteiden analysoinnissa. Tutkimuksen tarkoitus oli optimoida hedgerahaston tuotto sillä perusteella, miten uutisissa (Wall Street Journal) uutisoitiin ilmastomuutokseen liittyen. Tutkimuksessa pyrittiin myös luokittelemaan uutisten sisällöt positiivisiksi (sijoitusten arvoa nostaviksi) tai negatiivisiksi (sijoitusten arvoa vähentäviksi) sekä uutiset luokiteltiin myös koskemaan eri toimialoja (esim. energia, kulutustavarat yms.) (Enge et al. 2019). Todennäköisesti vastaavaa automaattisesti toimivaa pörssikauppaa toteutetaan jo nyt. Yle uutisoi vuonna 2017 vastaavantyyppisen tekoälyn käyttämisestä lyhytaikaisessa pörssikaupassa (Yle 2017b). Sosiaaliseen mediaan ja uutisiin perustuvissa suuraineistojen analysoinneissa olisi ehkä hyvä huomioida myös mahdolliset muut näkökulmat, kuten esim. kulttuurillinen tausta. Tietyistä vähäpätöisesti aiheista saatetaan uutisoida herkemmin mutta vastaavasti tärkeistä asioista ei välttämättä uutisoida yhtä laajasti. Tekniikka ja Talous -lehti uutisoi 2018 vastaavantyyllisistä automatisoiduista kaupankäynnistä pörssissä ja mainitaan riskeinä tilanteet, jossa algoritmit toimivat ohjelmallisesti oikein mutta algoritmien kaupankäynti saattaa aiheuttaa pörssikursseihin odottamattoman rajuja heilahduksia (Tekniikka ja Talous 2018). Hieman samantyyppisestä aiheesta on myös tehty diplomityö Tampereen Yliopistossa, jossa yritysten talousuutisiin liittyviä sisältöjä on analysoitu ja analyysien perusteella on pyritty ennakoimaan pörssikursseihin liittyviä epätavallisia tilanteita (Heikura 2018).

On huomattava, että vaikka suoranaisesti riskienhallintaan ja uhkien ennakkointiin liittyviä tutkimuksia ei ollut kovin monta, niin tarkastelussa oli ainoastaan suuraineistoihin perustuvat tutkimukset. Riskienhallintaa ja ongelmien ennakkointia käsitteleviä tutkimuksia on luonnollisesti enemmänkin, mutta näissä ei välttämättä ole käsitelty suuraineistoja vaan muita tutkimustapoja kuten esim. haastatteluja. Lisäksi myös vaikka maailman merenpinnan nousua ja tietyn eläinlajin uhanalaisuutta käsittelevä tutkimus voidaan luokitella myös riskinhallinnan alle, koska kyseisissä tutkimuksissa yritetään ennakoida asioita. Yksi syy riskienhallintaan liittyvien

tutkimusten vähyyteen voi olla se, että tämän tyyppistä tutkimusta ei välttämättä haluta julkaista avoimesti, koska riskienhallintaan liittyvä tieto voi olla osa yrityksen liiketoimintaa. Esimerkiksi ilmastonmuutoksen ilmiöiden vaikutusta analysoivassa sanomalehtien tekstianalyysissä (Enge et al. 2019) saatetaan liikkua alueella, josta yritykset eivät halua antaa toisilleen tietoja.

Global Environmental Change -lehdessä on tutkittu ilmastonmuutoksen vaikutusta turvapaikanhakijoiden siirtymiseen konfliktialueilta turvallisemmille alueille. Tutkimuksessa hyödynnettiin säätiöitä eri konfliktialueilta, 156 eri kohdemaalle tehtyjä turvapaikkahakemuksia vuosilta 2006-2015, sosioekonomisia (verkostot, väestömäärät, etniset ja poliittiset tiedot) ja maantieteellisiä aineistoja. Tutkimuksen mukaan ilmastonmuutoksen aiheuttamat ongelmat esim. maanviljelyksessä saattavat saada ihmiset muuttamaan kaupunkeihin, joissa työttömyys ja sosiaaliset ongelmat kasvavat. Mikäli hallinto ei pysty ratkaisemaan ongelmia, tilanne voi aiheuttaa erilaisia levottomuuksia ja aseellisia konflikteja, jotka puolestaan lisäävät turvapaikanhakijoiden määrää esim. naapurimaissa (Abel et al. 2019, Tekniikka ja Talous 2019). Tutkimus on hyvä esimerkki siitä, miten monia erilaisia suuraineistoja voidaan yhdistää ja analysoida. Ehkä yhtenä haasteena on suuraineistojen analysoinnin lisäksi myös se, millaisia suuraineistoja (mitä tietoja) tarvitaan. Ilmastopakolaiset olivat yksi suurimmista World Economic Forumin World Risk Report 2019 -julkaisussa määritellyistä maailman riskeistä (World Economic Forum 2019). Myös Yle on uutisoinut esim. vuonna 2009 ilmastonmuutoksen aiheuttamista ilmastopakolaisista (Yle 2009). Tämä tutkimus on myös hyvä esimerkki siitä, miten haastavaa on selvittää monimutkaisia syy ja seuraussuhteita ilmastonmuutokseen liittyen. Todennäköisesti suuraineistojen hyödyntäminen tämänkaltaisissa haasteissa on toimiva ratkaisu.

Käsitellyt riskienhallintaan liittyvät tutkimukset oli riskienhallinnan näkökulmasta melko tyyppillisiä ja tutkimuksiin liittyi tarkasti määriteltä riski, jota vastaan haluttiin suojautua. Data-Pop Alliancen julkaisussa vuodelta 2017 on käsitelty suuraineistoja ja riskienhallintaa laajemmin. Julkaisun mukaan suuraineistojen analysoinnilla pitäisi pyrkiä lisäämään yhteiskunnan kykyä (resilience) toipua erilaisia ilmastonmuutoksen aiheuttamista uhista (Data-Pop Alliance 2017). Tämä näkökulma tarkoittaisi sitä, että yksittäisen paikan maanjärityksen ennustamien sijaan pitäisi pyrkiä ennakoimaan laajempia ilmiöitä sekä tekemään analyyskejä, miten ilmiöihin voidaan reagoida ennen mahdollista katastrofia ja katastrofin tapahtumisen jälkeen. Tutkimuksen kannalta näin laaja ja epämääräinen tehtävänanto voi olla hyvin haastava ja todennäköistä on, että pelkkiä suuraineistoja analysoimalla ei näin laajoja tutkimuskysymyksiä voida ratkaista. Todennäköinen kehityssuunta on ehkä enempi jonkinlaiseen poikkitieteelliseen yhteistyöhön, jonka yhtenä osana on suuraineistoista tehtävät analyysit, joita erilaiset ja nimenomaan eri alojen asiantuntijat analysoivat riskienhallinnan näkökulmasta.

Edellä mainitun kaltaista, pidemmän ajan poikkitieteellistä ilmastonmuutokseen liittyvää tutkimusta tekee esim. IPCC (IPCC 2019) ja vastaavasti esim. valtiolliset ja kaupalliset toimijat todennäköisesti hyödyntävät erilaisia ilmastoon liittyviä suuraineistoja lyhyemmän ajan toiminnassa. Lyhyemmän ja pienemmän mittakaavan toiminnasta esimerkkinä (referoitujen tutkimusten lisäksi) on Espoon kaupungin suunnitelma arvioida kaikkien kaupungin päästöjensä ilmastovaikutukset (Yle 2018e).

4.3 Tyypillisiä aluespesifisiä tutkimuksia

Yksi tyypillinen osa-alue ilmastonmuutoksen vaikutuksia tutkittaessa on tutkimuksen keskittyminen johonkin rajattuun alueeseen (kuten esim. tiettyyn valtioon). Tyypillisesti tämän tyyllisissä tutkimuksissa on analysoitu jotain suuraineistoa ja on pyritty ymmärtämään esim. ilmastonmuutoksen vaikutusta alueelle. Tyypillinen rahoittaja tämän tyyllisissä tutkimuksissa on yleensä valtio tai suuryritys, joka tarkastelee vaikka tietyn maantieteellisen alueen investointimahdollisuuksia.

Tyypillinen ilmastonmuutosta käsittelevä tutkimus on esimerkiksi NASAn rahoittama Environmental Research Lettersin arktisen alueen jääpeitteen vuotuinen kasvaminen ja sulaminen. Aineistona tutkimuksessa on käytetty satelliittien keräämää tietoa (mikroaaltosäteily) arktisen alueen jääpeitteen käyttäytymisestä vuosina 1976-2016. Tutkimuksessa on analysoitu jääpeitteen jäätymistä ja sulamista ja havaittu, että kesän sulamiskausi on tilastollisesti merkittävästi pidentynyt tietyillä arktisilla alueilla. Tutkimuksen mukaan arktinen jääpeite on sulamassa (Bliss et al. 2019). Tutkimuksessa on jaettu aineisto osiin maantieteellisen sijainnin mukaan, tehty erilaisia analyysejä joita on visualisoitu, sekä analyyseistä on tehty johtopäätöksiä käyttäen erilaisia tilastollisia menetelmiä. Tutkimuksessa ei kuitenkaan ole otettu kantaa siihen, mistä arktisen alueen jäiden sulaminen tarkalleen ottaen johtuu ja todennäköisesti sulamissytyt liittyvät useisiin monimutkaisiin prosesseihin. Todennäköisesti mikäli jäätiköiden laajuutta mittaavien satelliittidatoiden lisäksi analysoitaisiin rinnalla myös muita vastaavia aineistoja, jolloin mahdollisesti saataisiin näiden aineistojen välisiä riippuvuussuhteita selville ja mahdollisesti myös uutta tietoa siitä, miten ilmastonmuutos etenee. Vastaavasti arktisilla alueilla toimivat kaupalliset ja valtiolliset toimijat voivat hyödyntää suuraineistojen analyysejä omissa liiketoiminnoissaan. Kuitenkin koska koko ilmaston toimintaa ei välttämättä tarkasti tunneta, nykyisillä havainnoilla ei välttämättä voida kovin tarkasti ennustaa tulevaa pidemmän aikavälin kehitystä. Vastaavasti yksittäisillä yrityksillä ei välttämättä ole resursseja kerätä omia suuraineistoja saati tehdä näihin liittyvää tutkimusta. Arktisen alueen jääpeitteen muutoksia käsittelevä tutkimus on hyvä malliesimerkki avointen suuraineistojen käyttämisestä ilmastonmuutokseen liittyvässä tutkimuksessa. Vastaavasti Yhdysvaltalaisessa tutkimuksessa on laskettu, että arktisten alueiden sulaminen nostaa merenpintaa ennakoitua enemmän (Bamber et al. 2019, Yle 2019k). Paikallinen ilmiö (arktisten alueiden jääpeitteen väheneminen) voi siis aiheuttaa globaaleja vaikutuksia, joita voi olla hankala ennakoita yhden tutkimuksen kautta.

Jääpeitteen sulaminen ja ilmaston muuttuminen vaikuttaa myös maapallolla kiertävään veteen. Journal of Water and Climate Change -lehdessä on julkaistu tutkimus, missä on mallinnettu maapallon veden hydrologista kiertoa Yhdysvaltojen Nevadassa sijaitsevalla Lehman Creek -vedenjakajalla. Tutkimuksessa käytettiin aineistona alueella sijaitsevan viiden sääaseman tietoja vuosilta 1981-2010. Havaintoja yhdisteltiin ja näistä luotiin matemaattinen malli, jolla myös tehtiin ennusteet vuoteen 2099 asti. Mallinnuksessa mallinnettiin myös vuosittaisiin vedenkiertoon vaikuttavia syklejä (Chen et al. 2019). Tämän tyyllisissä paikallisissa ennusteissa voi olla hankala mallintaa sitä, millaisia vaikutuksia erilaisilla globaaleilla ilmiöillä (kuten aiemmin mainittu tutkimus arktisen alueen lämpenemisestä) on tiettyyn rajattuun maantieteelliseen alueeseen. Lehman Creek -alueen suuraineistossa tarvittiin myös sääasemien mittaustulosten

yhdenmukaistamista ja puuttuvien tietojen huomioista. Ilmastomuutoksen kannalta haasteita voi tuoda myös äkilliset äärisäät, joiden ennustaminen matemaattisella mallilla on hyvin hankalaa. Yle uutisoi vuonna 2019 Mosambikin yllättävistä rankkasateista ja niitä seuranneista tulvista (Yle 2019l). Todennäköisesti Mosambikin alueellakin on tehty mallinnusta säähän ja vedenkiertoon liittyen, mutta erilaisten äärimmäisten sääilmiöiden ennustaminen voi olla haasteellista. Yle uutisoi jo 2011, että erilaisten äärisäiden ennustetaan tulevaisuudessa lisääntyvän (Yle 2011). Suomessa erityisiä tulvavaara-alueita on Ylen uutisen mukaan Pori ja Rovaniemi (Yle 2018f) ja vastaavanlaisia veteen liittyviä suuraineistoja kerätään ja analysoidaan Suomessakin mm. Ilmatieteen Laitoksen alaisessa Tulvatietokeskuksessa (Ilmatieteen Laitos 2019).

Environmetrics -lehdessä julkaistussa tutkimuksessa on mallinnettu Yhdysvaltojen ilmastoa käyttäen koneoppimista. Tutkimuksessa ilmastoon liittyvää dataa on annettu koneoppimisalgoritmin luokiteltavaksi ja saatuja luokittelutuloksia verrattiin olemassa oleviin (aiemmin tehtyihin ja virallisesti käytössä oleviin) alueiden säätyyppiluokituksiin. Käytettävissä olleessa suuraineistossa oli tietoja kolmen tuhannen sääaseman mittaustiedoista ympäri Yhdysvaltoja ajalta 1946-2015. Osa käytetyistä aineistoista käsitteli päivittäisi säätietoja, osan tarkkuus oli vuosi ja yksi aineisto käsitteli erilaisten äärisäiden esiintymisiä. Lisäksi tutkimuksessa tehtiin interaktiivinen verkossa toimiva datan visualisointi, jolla säätä voidaan luokitella reaaliaikaisesti (Sathiaraj et al. 2019). Vaikkei tutkimuksessa saatu mitään kovin suurta lopputulosta, tutkimus on hyvä esimerkki tyypillisestä suuraineistojen käyttämisestä data-analytiikan apuvälineitä hyödyntäen. Tutkimuksessa suuri datamäärä esitettiin visualisoinnin kautta ihmiselle helpommin ymmärrettävässä muodossa. Kuten tutkimuksessakin todettiin, ehkä suuraineistosta voitaisiin myös luoda malli, jolla tulevia säitä ennustettaisiin ja visualisoitaisiin. Käytetty data-aineisto oli kerätty maantieteellisesti rajatulta alueelta (Yhdysvallat) ja todennäköisesti muualla maailmassa tapahtuvat ilmastoilmiöt vaikuttavat myös Yhdysvaltojen tuleviin ilmasto-olosuhteisiin. Yle uutisoi vuonna 2018 hieman vastaavantyyllisestä ilmaston lämpenemiseen liittyvästä visualisoinnista, jossa maailman eri paikkojen lämpenemistä on visualisoitu NASAn avoimen datan aineiston avulla vuosilta 1880-2017 (Yle 2018g). Tyypillisesti ilmastomuutos saatetaan ymmärtää pelkkänä ilmaston lämpenemisenä, josta Environmetricsin tutkimus on hyvä esimerkki. Kuitenkin lämpenevä ilmasto vaikuttaa todennäköisesti hyvin monimutkaisella tavalla eri asioihin ja mm. yhteiskuntien elintapoihin ilmastomuutosta hillitessä. Sitran mukaan ilmastomuutoksen hillitseminen vaatisi Suomessa autoilun ja lihansyönnin vähentämistä (Yle 2019m). Tämänäyttelyisiä elämäntapojen ja yhteiskuntien muuttumisia voi olla hyvin hankala ennustaa tai mallintaa, toisin kuin vaikka lämpötilojen kehitystä.

Annals of Forest Science lehdessä julkaistussa tutkimuksessa on puolestaan tarkasteltu Saksassa tapahtuneita muutoksia. Käytetty suuraineisto on Saksan valtiollisten toimijoiden keräämää ilmastoon liittyvää tietoa vuosina 1961-2013. Ilmastoon liittyvä tieto (auriongon säteilyn määrä, lämpötilat, kosteuden haihtuminen, tuulen voimakkuus sekä sademäärä) on kerätty neliökilometrin kokoisista alueista. Saadun aineiston avulla on arvioitu myös mitattujen ilmastotietojen arvoja tulevaisuudessa vuoteen 2100 asti. Tutkimusta on ollut rajoittamassa mm. saksalaisia ydinenergiaan, rakentamiseen, luonnonsuojeluun ja ympäristöön liittyviä kansallisia toimijoita (Dietrich et al. 2019). Tutkimuksessa ei oteta kantaa siihen, mitä toimenpiteitä esim. Saksan valtion olisi syytä

tehdä ja todennäköisesti tutkimuksen tuloksia hyödynnetään eri virastoissa. Tutkimuksessa on tarkasteltu vain Saksan alueen tietoja, ja todennäköisesti esim. aiemmin mainittu arktisen alueen jääpeitteen sulaminen voi pahimmassa tapauksessa kiihdyttää ilmastonmuutosta tuntemattomilla tavoilla. Ylen uutisen mukaan arktisten alueiden jääpeitteen vähentyminen vaikuttaa myös globaaliin ilmastoon mm. lisäämällä erilaisia äärisäitä (Yle 2018b). Ehkä tällöin yksi tutkimuksen jatkokehityssuunnista voisi olla muiden tutkimusten yhdistäminen ja näistä (muista tutkimuksista) saatujen tietojen käyttäminen tulevaisuuden ennakkoinnissa. Kuitenkin, kuten big earth datalle on tyypillistä, erilaisten yhdistettyjen suuraineistojen monimutkaisuus nousisi todennäköisesti hyvin suureksi ja erilaisia monimutkaisia riippuvuuksia ja erilaisten ilmastojärjestelmien takaisinkytkentöjä olisi todennäköisesti hyvin hankala hahmottaa. Tutkimus on hyvä esimerkki siitä, miten suhteellisen yksinkertaisesta datasta (erilaisia säähavaintoja neliökilometrin alueelta vuosikymmenten aikana) voidaan johtaa ennusteita pitkälle tulevaisuuteen erilaisten valtiollisten toimijoiden käytettäväksi. Ylen uutisen mukaan tulevaisuudessa lämpenevä ilmasto lisää myös puiden ja muiden kasvien siitepölytuotantoa, joka puolestaan todennäköisesti lisää erilaisia allergioita (Yle 2008). Tämän tyyllisiä ilmiöitä voi olla hankala ennustaa pelkkää metsän ilmasto- ja sääolosuhteita mallintaessa.

Hieman samantapaisessa tutkimuksessa on mallinnettu ilmastonmuutoksen vaikutuksia Brasilian tuuli- ja aurinkoenergiaan liittyen. Tutkimuksessa on yhdistetty kolme globaalia ilmastomallia ja lopputuloksena on saatu tietoa siitä, missä maantieteellisessä osassa Brasiliassa tuulienergian ja aurinkoenergian määrä todennäköisesti tulevaisuudessa (2070-2099) kasvaa. Tutkimuksessa on mallinnettu tuulen, lämmön sekä auringon säteilyn määriä. Apuna on ollut IPCC:n ilmastomalli, jota käyttämällä on mallinnettu Brasilian alueen tulevaisuuden ilmasto-olosuhteita (de Jong et al. 2019). Tässä tutkimuksessa suuraineistojen käyttö tähtää liiketoimintamahdollisuuksien optimointiin sekä ilmastonmuutoksen vaikutusten hyödyntämiseen. Tutkimuksessa on yhdistetty liiketoiminnallista ja paikallista tietoa IPCC:n ilmastomalleihin, jolloin on saatu ennusteita tulevasta.

Applications in Plant Sciences -julkaisussa on tutkittu erilaisten sienien esiintymistä Pohjois- ja Keski-Euroopassa vuosina 1970-2010. Erityisesti on haluttu selvittää, miten ilmastonmuutos ja maankäyttö on vaikuttanut sienien esiintymiseen. Aineistona on ollut erilaisia museoiden ja harrastelijoiden havaintoja eri sienten esiintymisestä. Tutkimuksessa on lisäksi mallinnettu sienten elinolosuhteita käyttäen ilmastoon liittyviä tietoja, esiintyvien puulajien tietoja ja saasteita. Tällä mallilla on mallinnettu sienimääriä ja sienilajien monimuotisuutta. Aineistossa on ollut päiväkohtaisia tietoja ilmastosta sekä pidemmän aikavälin tietoja, ja näitä tietoja on yhdistelty. Mallinnuksen resoluutio on ollut eri asioille (esim. eri sienilajien esiintyvyys) 1 x 1 km. Tutkimuksen mukaan ilmaston lämpeneminen lisää erilaisten sienten esiintymistä luonnontilassa olevilla alueilla (Andrew et al. 2019). Tutkimus on hyvä esimerkki siitä, miten erilaisia avoimesti saatavia tietokantoja voidaan hyödyntää. Osa tietokannoista on ollut aiemmin mainitun Luonto-Liiton kevätseurannan tapaista harrastelijoiden keräämää tietoa ja erilaisia tietokantoja yhdistelemällä on saatu luotua uutta tietoa. Vaikka yksittäisen harrastelijan keräämät tiedot eivät vielä tuota kovin merkittävää tietoa, tiedon luotettavuutta ja uskottavuutta voidaan lisätä käyttämällä suuraineistoja. Ylen uutisen mukaan ilmastonmuutos sekoittaa jo nyt ekosysteemien toimintaa (Yle 2019n). Sienien

ja vastaavien yksittäisten lajien tutkimisessa hankaluutta todennäköisesti aiheuttaa luonnon monimutkaisuus. Vaikka yksittäisistä lajeista on käytettävissä monenlaista dataa, luonnon ekosysteemien erilaiset takaisinkytkennät ja riippuvuussuhteet voivat tehdä erilaisten ennusteiden tekemisestä hyvin haastavia. Vastaavasti laajempien ja useammista lähteistä peräisin olevien aineistojen käyttäminen voi tehdä data-analyysien teosta hyvin monimutkaista. Tivi-lehden uutisen mukaan Suomen valtiolle on erilaisia sähköisiä ja paperisia tietovarastoja, joiden hyödyntäminen ja koneluettavaan muotoon saattaminen vaatisi paljon työtä ja koordinoitua (Tivi 2019). Todennäköisesti hyvin samanlaisia ongelmia on myös erilaisten ilmastodataan liittyvien suuraineistojen kanssa.

Yhdysvaltalaisen Stanfordin yliopistossa on tutkittu ilmastomuutoksen vaikutusta eri maiden talouksiin. Tutkimuksessa on käyty läpi eri maiden bruttokansantuotteen kehitystä viidenkymmenen vuoden ajalta vuoteen 2010 asti. Tutkimuksen mukaan ilmaston lämpeneminen on hyödyttänyt pohjoisessa sijaitsevia maita ja vastaavasti lämpimän ilmaston maiden bruttokansantuote on tutkimuksen mukaan heikentynyt. Puolen vuosisadan aikana hyvinvointikuilu pohjoisen ja etelän välillä on kasvanut 25% (bruttokansantuotteen muutos) ilmastomuutoksen seurauksena (Diffenbaugh et al. 2019). On kuitenkin huomioitava, että tutkittu ajanjakso on melko lyhyt ja vaikka pohjoisessa sijaitsevat valtiot (kuten Suomi) ovat saaneet tutkimuksen mukaan etua ilmaston lämpenemisestä, niin ilmaston edelleen lämmetessä pitkän aikavälin vaikutukset voivat olla arvoitus. Talouden tunnuslukuja saatetaan myös mitata eri maissa ja vaikka bruttokansantuote lisääntyisikin, talouskasvun hyödyt eivät välttämättä jakaudu maissa samalla tavoin.

Ruuan tuotantoa ja erilaisten ilmastovaihtelujen vaikutusta tähän on tutkittu käyttäen julkisesti saatavia avoimen datan lähteitä. Ruuantuotanto eri maantieteellisillä aluilla on riippuvainen ilmaston muutoksista ja tuotannon turvaamiseksi on tehtävä pitkälle tulevaisuuteen meneviä ennusteita viljelymahdollisuuksien kehittymisestä. Aiemmin tämällyisissä ruokahuollon malleissa tarkasteltiin pelkästään taloudellisia, sosiaalisia ja ympäristöllisiä kriteerejä, mutta nykyään myös ilmaston muuttuminen pitkällä aikavälillä olisi huomioitava. Toiseksi, erilaisten epävarmuuksia vuoksi olisi syytä luoda ruokahuollon toimitusketju käyttäen yhden toimittajan sijaan useampia eri tavoin hajautettuja toimijoita. Ruokahuollon maantieteellisen sijainnin valitsemisen lisäksi pitäisi huomioida myös ilmastomuutoksen vaikutus ruuan hintoihin pitkällä aikavälillä. Tutkimuksessa ratkaisuna ongelmiin esitetään kolmivaiheista mallia (ilmaston vaikutuksen huomioiminen, ruuantuotannon riskien ja kasvuolosuhteiden huomiointi sekä toimitusketjuun liittyvien asioiden mallintaminen). Mallin luotaessa on käytetty erilaisia avoimen datan lähteitä liittyen ilmastoon ja maatalouteen (Srinivasana et al. 2019). Vaikka tutkimus kuulostaa ensilukemalta hyvin järkevältä, kannattaa muistaa, että tulevaisuuden ennustamiseen liittyy aina epävarmuutta ja mitä monimutkaisempi malli on kyseessä, sitä hankalampi mallia on todennäköisesti säätää. Ruuantuotannon toimitusketjuun voi vaikuttaa hyvin monet asiat, aina ruuan kuljettamisesta ruuan hintaan liittyen. Ilmastomuutoksen aiheuttamat äärisäät saattavat vaikuttaa globaaleihin merikuljetuksiin, joita todennäköisesti käytetään myös ruuan kuljettamiseen kuluttajille. Todennäköisesti aineistot liittyen eri maantieteellisten alueiden viljelymahdollisuuksiin sekä säähavaintoihin, ovat sinänsä luotettavia. Voi olla kuitenkin hankala ennustaa, miten ilmastomuutoksen eri vaikutukset vaikuttavat näihin. Ylen Uutisten mukaan esim. Suomessa kasvukausi on pidentynyt

ilmastonmuutoksen seurauksena, mutta erilaisten ilmastonmuutosten haittojen vaikutus on kuitenkin suurempi kuin hyötyjen (Yle 2018h). Todennäköisesti kuitenkin vastaavatyypistä avoimen lähteiden suuraineistojen käyttöä voidaan pitää hyvin järkevänä kehityssuuntana muillakin aloilla kuin tutkimuksessa mainitussa ruuantuotannon suunnittelussa.

Applied Energy -lehdessä on tarkasteltu ilmastonmuutoksen vaikutusta Los Angelesin sähkönkulutukseen. Tutkimuksessa oli yhdistetty tietoja alueen rakennuksista, mitatuista lämpötiloista (vuosina 1981-2000, mittausalueina kahden neliökilometrin alueet), alueiden asukasmääristä, arvioita käytössä olevista ilmastointilaitteista sekä alueen sähkönkulutuksesta. Näitä tietoja oli yhdistetty ja lopputuloksena oli luotu matemaattinen malli, millä sähkönkulutuksen muutosta voidaan arvioida tulevaisuudessa. Tutkimuksen mukaan ilmaston lämmitessä Los Angelesin sähkönkulutus tulee nousemaan, koska ilmastointilaitteiden käyttäminen lisääntyy. Tutkimuksen arvioita tulevaisuuden sähköntarpeesta voidaan käyttää kun alueen energiantarvetta suunnitellaan (Burillo et al. 2019). Tutkimus on hyvä esimerkki siitä, miten jo kerättyjä suuraineistoja voidaan yhdistellä ja yhdistelmää analysoimalla voidaan luoda uutta tietoa, jolle on selkeä loppukäyttäjä (alueen energialaitokset). Vaikka tutkimuksessa ei oltu puhuttu suuraineistoista (big data), niin tutkimuksessa oli silti suuraineistoja käytetty. Yle uutisoi kesäkuussa 2019 energiankäytön lisääntymisestä liittyen ilmastonmuutoksen aiheuttamiin ääri-ilmiöihin kuten kylmyyteen ja kuumuuteen (Yle 2019o). Energiankäytön ja ilmastonmuutoksen yhteys on hyvä esimerkki tutkimuksesta, jossa kahden asian (energiankäyttö ja ilmastonmuutos) suhdetta on tutkittu ja jonka tutkimustuloksilla on selkeä loppukäyttäjä (energiateollisuus ja energiantuotantoteollisuus).

Global Change Biology -lehdessä julkaistussa tutkimuksessa on tutkittu ilmastonmuutoksen vaikutusta karhujen elinympäristöön Espanjassa. Tutkimuksessa on yhdistelty eri tavoin kerättyjä suuraineistoja. Tutkimuksessa käytettiin vuosina 1985-2017 Espanjassa tehtyjä karhuhavaintoja. Lisäksi käytössä oli espanjalaisen valtiollisen toimijan keräämiä tietoja alueen metsistä (tarkkuus 1x1 neliökilometri) vuodelta 1998 sekä 2002-2003. Lisäksi käytössä oli tutkijoiden tekemä mallinnusaineisto kasvillisuuden (pääosin puita) esiintymisestä alueella. Näiden lisäksi käytettiin erilaisia ilmastoon (mm. sademäärät, kosteus, lämpötilat yms.) liittyviä mittaustietoja, jotka olivat kerätty vuosina 1960-1990. Lisäksi tutkijat olivat tehneet kaksi erilaista skenaarioita vuosille 2050 ja 2070, joissa ilmaston ennustettiin lämpenevän eri tavoin (kohtalainen lämpeneminen ja pessimistinen skenaario). Suuraineistoja oli analysoitu siten, että karhujen esiintymistä erilaisilla alueilla oli mallinnettu kerätystä datasta. Tämän jälkeen oli tehty ennusteet vuosille 2050 ja 2070 aiemman analyysin pohjalta koneoppimisalgoritmeja käyttäen. Mallinnuksen lopputuloksena saatiin 250 m x 250 m -tarkkuudella olevia karttoja karhujen elinolosuhteista alueella. Tutkimuksen mukaan ilmaston lämpeneminen muuttaa alueen kasvikantaa ja ilmastoa, joten karhujen elinolosuhteet alueella muuttuvat epäsuotuisiksi (Penteriani et al. 2019). Tutkimus on hyvä esimerkki siitä, miten erilaisia suuraineistoja voidaan yhdistellä, käyttää koneoppimisalgoritmeja mallintamaan ekosysteemiä ja mallinnuksen avulla voidaan luoda ennusteita tulevalla kehitykselle. Vastaavalla tavalla olisi todennäköisesti mahdollista mallintaa ilmastonmuutoksen vaikutusta muihinkin eläimiin ja kasveihin. Karhupopulaation väheneminen voi aiheuttaa myös muutoksia alueen ekosysteemiin ja

tällä voi olla vaikutuksia alueen kasveihin ja eläimiin. Vähenemisen vaikutusta voi olla hankala mallintaa, mikäli alueella on aina ollut karhuja eikä ole tiedossa, miten karhupopulaation väheneminen vaikuttaa ympäröivään ekosysteemiin. Yle uutisoi kesällä 2019 *Nature ecology & evolution* -lehdessä julkaistusta tutkimuksesta, minkä mukaan viimeisten 250 vuoden aikana 571 kasvilajia on kuollut sukupuuttoon ja kasvilajien sukupuuttoon kuoleminen nopeus on kiihtynyt samassa ajassa viisisataakertaiseksi (Yle 2019p). Vastaaventyypiset, ilmastonmuutoksen aiheuttamat muutokset ekosysteemeihin sekä näistä johtuvat eläin- ja kasvikantojen muutokset ovat hyvin tyypillisiä ilmastonmuutokseen liittyviä uutisia. Erilaisia ekosysteemejä mallinnettaessa on huomioitava myös lajien leviäminen. Ylen uutinen kesältä 2019 kertoo vieraslajina levinneestä mustatäplätokosta, jonka leviäminen Perämerelle on uhka paikallisille lajeille (Yle 2019q). Tulevaisuuden ekosysteemejä mallinnettaessa vieraslajien leviäminen vaatii todennäköisesti uusien suuraineistojen keräämistä, jossa on esim. havaintoja vieraslajien leviämistä uusille alueille. Kaiken kaikkiaan ekosysteemien mallintaminen on todennäköisesti hyvin haastavaa, mutta erilaisten suuraineistojen yhdistämisellä pystytään saamaan jonkinlainen käsitys siitä, mitä ekosysteemeissä tapahtuu ilmaston lämmetessä. Kuitenkin on huomattava, ettei ilmaston lämpenemisestä tulevaisuudessa ole tarkkoja tietoja, joten erilaiset tulevaisuusskenaariot ekosysteemeihin ja muihin vastaaviin liittyen ovat aina lähinnä erilaisia spekulatioita. *Global Change Biology* -julkaisussa oli oivallisesti eroteltu arvioidun lämpenemisen skenaario sekä kiihtyneen ilmastonmuutoksen aiheuttama skenaario. *Ecological Applications* -lehdessä on julkaistu samantyyppinen tutkimus liittyen ekosysteemien muuttumiseen ja eläinlajien elinolosuhteisiin ilmastonmuutoksessa. Tässä tutkimuksessa on käytetty myös suuraineistoja mutta suuraineistossa olevia puutteita ja tulevaisuusskennustaiden epävarmuuksia on paikattu käyttäen apuna tutkittavien eläinlajien ja ekosysteemien asiantuntijoita, jotka ovat olleet apuna tehtäessä mallinnusta (Reside et al. 2019).

Nature-lehdessä julkaistussa tutkimuksessa on analysoitu 80 000 korkearesoluutioista satelliittikuvaa. Kuva-aineiston lisäksi käytössä oli kymmenen erilaista muuttujaa, jotka kuvasivat mm. maaperää, ilmastomalleja, pinnanmuotoja sekä tietoja paikallisista puulajeista. Kuvia oli analysoitu Google Earth Engine -sivuston ja koneoppimismenetelmien avulla. Tutkimuksessa etsittiin satelliittikuvista paikkoja, joihin voitaisiin vielä istuttaa puita ilman että istutetut puut haittaisivat maanviljelystä tai ihmisten elämää. Lisäksi haluttiin selvittää, millaisen määrän hiilidioksidia nämä istuttavat puut voisivat sitoa. Tutkimuksen perusteella maapallolle voitaisiin istuttaa niin paljon puita, että puiden kyky sitoa ilmakehän hiilidioksidia olisi erittäin suuri. Sidottavan hiilidioksidin määrä on kaksi kolmasosaa siitä määrästä, mitä ihmiskunnan toiminnan seurauksena on vapautunut ilmakehään teollisen vallankumouksen jälkeen. Tutkimuksen mukaan tämä (puiden istuttaminen) olisi ylivoimaisesti tehokkain tapa sitoa hiilidioksidia ilmakehästä (Bastin et al. 2019, Yle 2019r, Tekniikan Maailma 2019a). Tutkimukseen liittyy nettisivusto, jossa puiden istutuspaikoista esitetään visualisointi (Crowtherlabs 2019). Tutkimus ilmestyi samaan aikaan sekä *Nature*-lehdessä ja samana päivänä myös Ylen uutisissa. Vaikka tutkimus ei varsinaisesti ollut aluespesifinen, tutkimuksesta voidaan myös saada aluespesifistä tietoa. Verrattuna muihin tässä dokumenteissa käsiteltyihin tutkimuksiin, tämä *Nature* -lehden tutkimus on mahdollisesti kunnianhimoisin, sillä tutkimuksen mukaan pelkkä puiden istuttaminen olisi ylivoimaisesti tehokkainta ilmastonmuutoksen torjuntaa. Aiemmin mainituissa

National Geographic -lehdessä mainittiin kuitenkin, että pelkkää data-analyysiä käyttämällä osa tiedoista saattaa jäädä piiloon (Boas et al. 2019). Ehkä Nature-lehden tutkimukseen kannattaisi yhdistellä muista tutkimuksista ja muista suuraineistoista saatuja tietoja. Tällöin voitaisiin selvittää esim. että miksei Nature-lehden tutkimuksessa määritetyillä puuttomilla alueilla (jonne puita on tarkoitus istuttaa) kasva vielä puita. Ehkä National Geographic -lehden tyyliä paikallisia ihmisiä haastatteleamalla olisi saatu tietoa tiettyjen alueiden metsätaloudesta ja kulttuurista, jolloin asiasta olisi ehkä saatu kattavampi kokonaiskuva. Nature-lehden tutkimuksessa globaali-ilmastomalli pitäisi myös muuntaa paikallisiksi malleiksi (”downscaling”, Schneider et al. 2019).

Nature -lehdessä julkaistussa tutkimuksessa oli hyödynnetty ilmastoon liittyviä aineistoja viimeisten kahden tuhannen vuoden ajalta. Tutkimuksessa oli hyödynnetty kansainvälisen PAGES-tutkimuskonsortion keräämiä suuraineistoja. Käytetyissä suuraineistoissa oli tietoa mm. lämpötiloista, jäätiköistä, koralleista ja puiden vuosirenkaista. Kaikki suuraineistot eivät olleet kahden tuhannen vuoden ajalta, vaan osa aineistoista oli lyhyemmältä, esim. sadan vuoden ajalta. Tutkimuksessa käytetty ohjelmakoodi sekä käytetyt suuraineistot olivat myös vapaasti saatavilla. Tutkimuksessa oli tarkasteltu ilmaston muuttumista kahden vuosituhannen aikana. Tutkimuksen mukaan ilmaston muutokset vaihtelivat maantieteellisen sijainnin mukaan. Suuraineistojen analysoinnin perusteella ilmastomuutos oli kiihtynyt merkittävästä viimeisen sadan vuoden aikana ja ilmaston lämpeneminen on tällä hetkellä erittäin nopeaa (Neukom et al. 2019, Tekniikan Maaailma 2019b, PAGES 2019). Tutkimus on hyvä esimerkki siitä, miten eri lähteistä (jäätiköt, puiden vuosirenkaat yms.) kerättyä dataa on pystytty hyödyntämään ja luomaan pitkiä aikavälejä käsittäviä suuraineistoja. Lopputuloksena on saatu hyvin vakuuttavaa tietoa, jonka luotettavuus on todennäköisesti hyvin korkea. Ylen Uutisissa erilaiset korkeat lämpötilat ovat olleet kuluneina kesinä toistuva teema (esim. Yle 2019s). Tutkimus on hyvin vakuuttava esimerkki erilaisten vapaasti saatavilla olevien suuraineistojen käyttämisestä. Kun analyysien pohjana olevat suuraineistot ovat vapaasti kenen tahansa saatavilla ja analysoitavissa samoilla tutkijoiden käyttämällä ohjelmakoodilla, niin lopputulosta on hyvin hankala kiistää. Lisäksi mahdollisia virheitä ja epätasaisuuksia on helpompi löytää, mikäli analyysien tekeminen on mahdollista kelle tahansa. The Guardian lehdessä olevassa kyseistä tutkimusta käsittelevässä artikkelissa todetaan, että tämä tutkimus toivottavasti osoittaa lopullisesti, että ilmaston lämpeneminen ei ole luonnollista ja johtuu ihmisen vaikutuksesta (The Guardian 2019). Erilaiset yksittäiset ja aluespesifiset tutkimukset on mahdollista kiistää ja ehkä selittää jollain luonnollisella syyllä. Vastaavasti useampi suuraineistoja käyttävät tutkimukset voivat olla hankalampia kiistää, eteenkin jos samoja analyysejä pystyy kuka tahansa tekemään. Ehkä tutkimuksen tulokset olisivat vieläkin vakuuttavammat, mikäli tutkimukseen olisi liittynyt joku interaktiivinen verkkosivusto, josta kuka tahansa (muukin, kuin data-analyysejä tekevä) voisi jossain mittakaavassa tehdä samoja data-analyysejä ja ehkä muuttaa jotain yksinkertaisia parametreja. Lisäksi tutkimukseen voitaisiin ehkä tulevaisuudessa yhdistellä mittaustietojen lisäksi esim. sosio-ekonomisia tietoja ja muiden (kuin luonnontieteiden) tuottamia suuraineistoja.

Aluespesifiset tutkimukset ovat tyypillisiä ilmastomuutosta tutkiva ja suuraineistoja käyttäviä tutkimuksia. Osassa tutkimuksia suuraineiston lisäksi käytettiin myös muita tutkimustapoja, esim. alueen väestön haastatteluja. Muut tutkimustavat toivat

tyypillisesti lisää tietoa ja ilman esim. haastatteluja, data-analyysistä saadut tiedot olisivat voineet jäädä puutteellisiksi.

Vastaavan tyyllisiä tutkimuksia on tehty pidemmän aikaa ja digitaalisten suuraineistojen käyttäminen ja aineistojen saatavuus ovat mahdollistaneet yhä laajempia ja hajautetumpia tutkimuksia.

5 JOHTOPÄÄTÖKSET

Tässä kappaleessa pohditaan referoituja tutkimuksia ja uutisartikkeleita tutkimuskysymyksen sekä alikysymysten kannalta.

5.1 Referoidut tutkimukset ja 4V-malli

Referoituja suuraineistoja voidaan tarkastella myös 4V-mallin kautta. 4V-mallissa on suuraineistoille oli määritelty neljä erilaista tuntomerkkiä. Referoidut tutkimukset voidaan luokitella siten, mitä 4V-mallin tuntomerkkiä kyseinen tutkimus vastaa. Luonnollisesti erilaiset mittausaineistot sisältävät paljon dataa, jossa voi olla virheitä ja epämääräisyyksiä sekä erilaisia trendejä. Referoitujen tutkimusten suuraineistot soveltuvat luonnollisesti osaltaan jokaiseen 4V-mallin kategoriaan jossain mittakaavassa.

Koska tarkasteltavana on nimenomaan ilmastonmuutokseen liittyvä tutkimus, on määriteltävä, miten 4V-mallia sovelletaan nimenomaan tässä kirjallisuustutkimuksessa. Alla olevassa taulukossa on yksi luokittelu, jonka luokitteluperusteet on esitetty alla.

- **Volume:** Tutkimuksessa käytetty suuraineiston näkökulmasta poikkeuksellisen suurta aineistoa. Tämä näkyy erityisesti siten, että suuraineiston analysointi vaatii erittäin raskasta laskentatehoa.
- **Variety:** Tutkimuksen erikoispiirteenä on hyvin erilaisten suuraineistojen yhdistely, joissa toisistaan huomattavasti poikkeavia aineistoja on yhdistelty.
- **Velocity:** Tutkimukset, joissa käytetty suuraineisto saattaa sisältää paljon dataa, joka muuttuu hyvin nopeasti. Esimerkkinä on sosiaalisesta mediasta saatavaa dataa, jonka analyysien tekeminen ja hyödyntäminen vaatii jonkinlaista reaaliaikaisuutta.
- **Veracity:** Tutkimukset, joissa käytetty suuraineisto sisältää paljon epävarmaa ja vaikeasti tulkittavaa dataa.

Referoitu tutkimus	Suuraineiston kuvaus	4V Volume	4V Variety	4V Velocity	4V Veracity
Boas et al. 2019	Erilaisia paikkatietoja	x		x	
Brönnimann et al. 2019	Säädata, äärisäädata	x			
Muszynski et al. 2019	Ilmakehä-/ilmastodata	x			
Michetti et al. 2019	Säädata,	x	x		x

Referoitu tutkimus	Suuraineiston kuvaus	4V Volume	4V Variety	4V Velocity	4V Veracity
	sosioekonominen data, maanviljelysdata				
Shukla et al. 2019	Säädata, maanviljelysdata	x			
Tiampo et al. 2019	Maanjärstysdata	x			
Engle et al. 2019	Sanomalehtien uutissisällöt			x	x
Abel et al. 2019	Säädata, sosioekonominen data, maantieteellinen data, yms.	x	x		
Bliss et al. 2019	Jääpeitedata, lämpötiladata	x			
Bamber et al. 2019	Säädata, jääpeitedata	x			
Chen et al. 2019	Vesidata, säädata	x			
Sathiaraj et al. 2019	Säädata	x			
Dietrich et al. 2019	Säädata	x			
de Jong et al. 2019	Aurinko-, vesivoima- ja tuulienergiadata	x			x
Li et al. 2019	Energiankäyttöön ja tuottamiseen liittyvää dataa				x
Andrew et al. 2019	Säädata, sienidata, kasvidata, maaperädata	x	x		x
Diffenbaugh et al. 2019	Bruttokansantuote, säädata	x	x		
Srinivasana et al. 2019	Säädata, säädatan ja viljelydatan tilastollinen jakauma, riskidata	x	x	x	x
Burillo et al. 2019	Säädata, rakennuksiin liittyvä data, sähkönkulutusdata	x	x	x	x
Penteriani et al. 2019	Karhidata, kasvidata, säädata	x			x
Reside et al. 2019	Kasvidata, eläindata, säädata	x	x	x	x
Bastin et al. 2019	Korkearesoluutioisia satelliittikuvia	x			

Referoitu tutkimus	Suuraineiston kuvaus	4V Volume	4V Variety	4V Velocity	4V Veracity
Neukom et al. 2019	Useita suuraineistoja liittyen säätietoihin, puiden vuosirenkaisiin yms.	x	x		x

Taulukko 3. Referoidut tutkimukset 4V-mallissa.

Vaikka 4V-mallin mukainen luokittelu on vain suuntaa antava, ja perustuu pitkälti luokittelijan itsensä näkemykseen, saadaan luokittelusta kuitenkin joitain karkeita suuntaviivoja. Tyypillisesti ilmastomuutosta käsittelevien tutkimusten suuraineistot näyttäisivät olevan erityisen paljon laskentatehoa vaativia (4V-mallin "volume"). Kuitenkin tarkasteltavissa tutkimuksissa oli selkeästi muutama tutkimus, missä oli yhdistelty erilaisia suuraineistoja (Srinivasana et al. 2019, Burillo et al. 2019, Reside et al. 2019, Neukom et al. 2019). On myös huomattava, että vaikka sosiaalisen median merkitystä korostetaan nykyisessä yhteiskunnassa, sosiaalista mediaa tai vastaavia aineistoja hyödyntäviä tutkimuksia oli vähän (lähinnä ainoastaan Engle et al. 2019).

Vaikka suuraineistot ovat eräänlainen trenditermi, suuraineistojen käyttäminen näyttäisi olevan enempi 4V-mallin "volume" -kategoriaan kuulumista. Vaikka tutkimuksissa puhuttiin nimenomaan suuraineistojen käyttämisestä ja käytettiin yleensä "big data" -termiä, termin käyttö kuvastaa ehkä enempi sitä, että tutkimuksen tekijät haluavat käyttää trenditermejä. Vastaavia suuraineistotutkimuksia on aiemminkin tehty, joissa ei välttämättä käytetty trenditermejä (esim. Mäkinen 2019 on käytännössä suuraineistoa hyödyntävä ilmastomuutostutkimus ilman trenditermejä).

Sen sijaan referoiduissa tutkimuksissa on nähtävissä viitteitä siitä, että kenties tulevaisuudessa suuraineistoja yhdistellään yhä enemmän muihin suur- ja pienaineistoihin, jolloin saadaan todennäköisesti huomattavasti kattavampaa tietoa.

5.2 Suuraineistojen hyödyntämisen malli

Referoitujen tutkimusten perusteella tyypillinen suuraineistojen optimaalinen käyttötilanne voisi olla esimerkiksi seuraavan kuvitteellisen esimerkin mukainen. Esimerkkinä on käytetty kuvitteellista jäätikön tutkimusta.

- Tutkittavasta jäätiköstä on kerätty dataa (esim. satelliittikuvia) pitkältä aikaväliltä (kymmeniä vuosia). Kerättyä suuraineistoa analysoidaan ja saadaan selville se, miten jäätikkö on vuosikymmenten aikana muuttunut sekä miten mahdollisesti jäätikkö tulevaisuudessa muuttuu.
- Jäätikködataan yhdistetään muita suuraineistoja, esim. kerättyjä kasvinäytteitä ja eläinhavaintoja pitkältä ajalta (kymmeniä vuosia). Jäätikön sulamisen ja kasvamisen yhteyttä alueen eläimiin ja kasveihin pystytään selvittämään aineistoja yhdistämällä. Yhdistetystä suuraineistosta saadaan kattavampia ennusteita sille, mitä tulevaisuudessa mahdollisesti tapahtuu alueen kasveille ja eläimille. Lisäksi voidaan arvioida, minkälaisia vuorovaikutuksia eri alueille

aiheutuu jäätiköllä tapahtuvista muutoksista.

- Muita suuraineistoja, esim. globaaleja sää tietoja pitkältä ajalta voidaan yhdistää edellä mainittuihin aineistoihin, ja saada käsitys siitä, miten erilaiset globaalit ilmiöt (esim. tuulet, globaali ilmasto, erilaiset päästöt) ovat yhteydessä paikallisesti tutkittuun jäätikköön.
- Edellä mainittuihin suuraineistoihin yhdistetään erilaisia sosio-ekonomisia tietoja, joiden avulla voidaan havaita jäätikön muutosten riippuvuussuhteita alueen ihmisten elämään sekä mahdollisesti ennakoita tulevaa kehitystä.
- Suuraineistoanalyysien tulkintojen apuna voidaan käyttää esim. asiantuntija-arvioita, haastatteluja tai muuta perinteisemmän ja pienimuotoisemman tutkimuksen työkaluja.
- Tehtyjä malleja ja analyysejä sekä tulevaisuusskenaarioita voidaan eri tavoin visualisoida ja käyttää tietoa hyväksi erilaisessa päätöksenteossa, ennakoinnissa ja riskienhallinnassa.

Käytännössä tällaisen todennäköisesti yhdenkin suuraineiston analysointi ja tulkinta voi olla haastavaa, mutta periaatteessa suuraineistoja yhdistelemällä on mahdollista luoda uutta tietoa, jota on hyvin vaikeata saada muulla tavoin. Kun tutkittavat ilmiöt ovat hyvin monimutkaisia, ilmiöillä on erilaisia mahdollisesti tuntemattomiakin riippuvuussuhteita, suuraineistojen käyttäminen, analysointi ja yhdistäminen on todennäköisesti järkevä lähestymistapa.

Kuvitteellinen esimerkki on jonkinlainen kooste siitä, miten suuraineistoja voitaisiin hyödyntää ilmastonmuutokseen liittyvässä tutkimuksessa. Malliin on koottu ideoita ja totutustapoja referoiduista tutkimuksista. Mallin numerointi on vain suuntaa antava, ja todennäköisesti vaiheiden tarkempi järjestys riippuisi tutkittavasta ilmiöstä ja käytettävissä olevista suur- ja pienaineistoista.

Mallin (tutkimuksen) vaihe	Esimerkki
1. Suuraineiston muuttaminen digitaalseksi	Puun vuosirengastiedon muuttaminen numeeriseen (ajalliseen) tietoon
2. Analysoitavan suuraineiston koostaminen	Puun vuosirengastiedon ja mitattujen lämpötilojen sekä jääpeitetietojen yhdistäminen
3. Suuraineiston analysointi ja matemaattisten mallien luominen	Matemaattinen malli siitä, miten lämpötila vaikuttaa puiden kasvuun ja jäätiköihin
4. Asiantuntijoiden käyttämien analyysissä apuna	Matemaattisen mallin tulkinta ja parametrien säätäminen
5. Pienaineistojen ja muiden suuraineistoanalyysien yhdistäminen	Paikallisten asukkaiden haastattelut, merien lämpötilamallinen hyödyntäminen
6. Tulosten visualisointi	Tulosten esittäminen esim. kartalla
7. Tulosten hyödyntäminen paikallisessa päätöksenteossa	Paikallisyhteisöjen elinkeinojen tulevaisuuden suunnittelu

Taulukko 4. Hahmotelma suuraineistojen hyödyntämiselle ilmastonmuutokseen liittyvässä tutkimuksessa.

On myös huomattava, että vaikka suuraineistojen analysoimalla saataisiin tietoa ilmastonmuutoksesta ja sen mahdollisista tulevaisuuden vaikutuksista, tarvittaviin

toimenpiteisiin saattaa vaikuttaa esim. eri valtioiden politiikka sekä monimutkaiset taloudelliset intressit.

5.3 Johtopäätösten yhteenvedoa ja analysointia

Käsitellyissä, vuonna 2019 julkaistuissa tutkimuksissa (34 valittua tutkimusta) käsiteltiin erilaisia ilmastomuutokseen liittyviä suuraineistoja ja niistä tehtäviä analyysejä. Tutkimusten perusteella voidaan tehdä seuraavansuuntaisia johtopäätöksiä.

On huomattavat, että vuonna 2019 julkaistujen ilmastomuutosta jollain tavalla käsittelevien tutkimusten lukumäärä on arviolta yli 100 000 (karkea ennuste perustuu Google Scholarin "climate change" -hakusanalla tehtyyn hakuun vuoden 2018 tutkimuksista). Tässä tutkimuksessa on tarkasteltu tarkemmin ainoastaan 34 tarkoin valittua tutkimusta, joten tämän lopputyön yhteenvedo koskee ainoastaan näitä tutkimuksia, niiden näkymistä Ylen uutisoinnissa ja suuraineistojen hyödyntämistä juuri näissä valituissa tutkimuksissa.

Suuraineistojen datamäärät ovat koko ajan kasvamassa kun erilaista dataa saadaan eri lähteistä koko ajan yhä enempi. Datamäärien kasvaminen luo uudenlaisia ongelmia, jotka liittyvät data-analyysien monimutkaistumiseen, datan tulkittavuuteen ja datasta tehtyjen johtopäätösten analysointiin. Datan analysointi vaatii yhä monimutkaisempia keinoja ja yhä pidemmälle vietyä osaamista eri alojen ihmisiltä. Suuraineistojen tulkinta vaatii myös täsmätietoa eri aloilta ilmastomuutoksesta psykologiaan. Suuraineistojen ja analyysien monimutkaistumisessa tarvitaan myös yhä enempi erilaisia yksinkertaisempia malleja ja datan visualisointeja, jotta analyyseistä saatava tieto saadaan paremmin ymmärrettävään muotoon. Ilmastomuutoksessa ongelmia aiheuttaa myös se, että tutkittavat ilmiöt saattavat muuttua hitaasti ajan kuluessa ja suuraineistoja täytyy kerätä ja analysoida pidemmältä aikaväliltä. Lisäksi ilmastoon ja ilmastomuutoksen tutkimukseen voi liittyä sellaisia asioita ja ilmiötä, joiden olemassaolosta, vaikutuksista ja vuorovaikutussuhteista ei ole vielä tietoa. Näitä tuntemattomia asioita voidaan myös osaltaan saada selville erilaisia suuraineistoja analysoimalla.

Alla olevassa taulukossa on esitetty karkealla tasolla tarkastellut vuonna 2019 julkaistut erilaisia suuraineistoja käyttävät referoidut tutkimukset.

Tutkimustyyppi	Referoitu tutkimus
Yleisiä tutkimuksia suuraineistojen käyttämisessä ilmastomuutokseen liittyen	Tarvitaan suuraineistoja, datan ymmärtämistä ja ennustamista (Hassani et al. 2019)
	Earth Observation Data -termi, datan ymmärtäminen keskeistä (Sudmanns et al. 2019)
	Big Earth Data -termi, systeemiarkkitehtuurimalli (Yang et al. 2019)
	Suuraineistojen ja paikallisten ihmisten haastattelujen yhdistäminen (Boas et al. 2019)

Tutkimustyyppi	Referoitu tutkimus
	Suuraineistojen käyttäjäryhmät ja käytön motiivit (Lassinantti et al. 2019)
	Suuraineistojen analysoinnilla lisäarvoa kestävä kehityksen firmoille (Singh et al. 2019)
	Kaupungeista kerättyjen suuraineistojen hyödyntäminen (Creutzig et al. 2019)
	Smart Cities -termi, kaupungeista kerättyjen suuraineistojen mahdollisuudet (Allama et al. 2019)
	Biodiversiteettidatan erikoispiirteitä (König et al. 2019)
	Koneoppiminen ja data-analyysien haasteet (Reichstein et al. 2019)
	Suuraineistojen käyttäminen energiateollisuudessa (Li et al. 2019)
	Globaalien datavisualisointien muuttaminen paikalliseksi toiminnaksi (Schneider et al. 2019)
Suuraineistojen käyttäminen riskienhallinnassa	Äärisäiden ennustaminen (Brönnimann et al. 2019, Muszynski et al. 2019).
	Metsäpalojen esiintymistodennäköisyydet Italian eri osissa (Michetti et al. 2019).
	Ilmastonmuutoksen riskit Himalajan alueen maanviljelykselle (Shukla et al. 2019).
	Maanjäristysten esiintymistodennäköisyydet Kanadan Montrealissa (Tiampo et al. 2019).
	Wall Street Journal lehden ilmastonmuutosaiheisen sana-analyysin käyttö sijoittamisessa (Engle et al. 2019).
	Ilmastonmuutoksen vaikutus turvapaikanhakijoiden määrään (Abel et al. 2019).
Aluespesifisiä tutkimuksia	Arktisen alueen jääpeitteen sulaminen (Bliss et al. 2019).
	Ilmaston lämpenemisen vaikutus jokisuistossa Yhdysvalloissa (Bamber et al. 2019, Chen et al. 2019).
	Yhdysvaltojen ilmasto-olosuhteiden luokittelu (Sathiaraj et al. 2019).
	Saksan metsien kasvuolosuhteet (Dietrich et al. 2019).
	Brasilian aurinkosäteilyn jakautuminen alueittain tulevaisuudessa (de Jong et al. 2019).
	Suuraineistojen käyttäminen energia-alalla (Li et al. 2019).
	Pohjois- ja Keski-Euroopan sienilajien elinolosuhteet (Andrew et al. 2019).
	Ilmastonmuutoksen vaikutus taloudelliseen menestykseen (Diftenbaugh et al. 2019).
	Ruokahuollon suunnittelu (Srinivasana et al. 2019).
	Ilmastonmuutoksen vaikutus Los Angelesin sähkökulutukseen (Burillo et al. 2019).

Tutkimustyyppi	Referoitu tutkimus
	Ilmastonmuutoksen vaikutus karhujen elinolosuhteisiin Espanjassa (Penteriani et al. 2019).
	Asiantuntijoiden käyttäminen suuraineistojen rinnalla (Reside et al. 2019).
	Satelliittikuvien käyttäminen puiden kasvatuspaikkojen etsintään (Bastin et al. 2019).
	Useiden suuraineistojen yhdistäminen (Neukom et al. 2019).

Taulukko 5. Yhteenveto referoiduista tutkimuksista.

Alla olevassa kuvassa on esitetty karkealla tasolla referoitujen tutkimusten käyttämät erilaiset suuraineistot sekä muut aineistot.

Referoitu tutkimus	Vapaamuotoinen kuvaus käytetystä suuraineistosta	Käytetyt ei-suuraineistot
Boas et al. 2019	Erilaisia paikkatietoja	Paikallisten haastattelut
Brönnimann et al. 2019	Säädata, äärisäädata	
Muszynski et al. 2019	Ilmakehä-/ilmastodata	
Michetti et al. 2019	Säädata, sosioekonominen data, maanviljelysdata	
Shukla et al. 2019	Säädata, maanviljelysdata	Maanviljelijöiden haastattelut
Tiampo et al. 2019	Maanjäristysdata	
Engle et al. 2019	Sanomalehtien uutissisällöt	
Abel et al. 2019	Säädata, sosioekonominen data, maantieteellinen data, yms.	Konfliktien tietoja
Bliss et al. 2019	Jääpeitedata, lämpötiladata	
Bamber et al. 2019	Säädata, jääpeitedata	
Chen et al. 2019	Vesidata, säädata	
Sathiaraj et al. 2019	Säädata	
Dietrich et al. 2019	Säädata	
de Jong et al. 2019	Aurinko-, vesivoima- ja tuulienergiadata	
Li et al. 2019	Energiankäyttöön ja tuottamiseen liittyvää dataa	
Andrew et al. 2019	Säädata, sienidata, kasvidata, maaperädata	
Diffenbaugh et al. 2019	Bruttokansantuote, säädata	
Srinivasana et al. 2019	Säädata, säädatan ja viljelydatan tilastollinen jakauma, riskidata	
Burillo et al. 2019	Säädata, rakennuksiin liittyvä data, sähkönkulutusdata	
Penteriani et al. 2019	Karhudata, kasvidata, säädata	

Referoitu tutkimus	Vapaamuotoinen kuvaus käytetystä suuraineistosta	Käytetyt ei-suuraineistot
Reside et al. 2019	Kasvidata, eläindata, säädata	Asiantuntija-arviot
Bastin et al. 2019	Korkearesoluutioisia satelliittikuvia	
Neukom et al. 2019	Useita suuraineistoja liittyen sää tietoihin, puiden vuosirenkaisiin yms.	

Taulukko 6. Yhteenvedo referoitujen tutkimusten käyttämisestä suuraineistoista.

Tyypillinen suuraineistoja käyttävä tutkimus liittyy joko datamassojen käyttämiseen yleisesti tai johonkin alueellisesti ja tutkimuksellisesti rajattuun asiaan (esim. jonkun lajien leviämisen ennustaminen). Erilaiset kaupallisiin tarkoituksiin liittyvät julkaistut tutkimukset käsittelevät pääosin pitkän aikavälin yritysstrategioita liittyen esim. ruuantuotannon ja energian tuotantoon. Suuraineistoja käyttävät erityisesti riskinhallintaan ja ongelmien ennakointiin liittyvät tutkimukset olivat harvinaisempia. Tämä saattaa johtua siitä, että tämän tyylliset tutkimukset voidaan helposti mieltää liikesalaisuuksiksi eikä näitä haluta julkaista kaikille. Liiketoimintaan liittyvissä tutkimuksissa nousi esille erityisesti pohdinta siitä, miten yritysten liiketoiminta saadaan toimiaan kannattavasti ilmastonmuutoksen häiritsemässä tulevaisuudessakin.

Suuraineistojen käyttäminen monimutkaisissa tarkoituksissa (esim. ilmastonmuutoksen ennakointi) vaatii myös hyvin paljon matemaattista ja liiketoiminnallista osaamista. Yksi tulevaisuuden kehityssuunta voi olla suuraineistoista tehtävien mallien ja analyysien siirtymisen omaksi liiketoiminnakseen. Yrityksillä itsellään ei välttämättä ole resursseja kehittää monimutkaisia analyysimalleja eikä tämä välttämättä ole yritysten ydinliiketoimintaakaan. Tästä syystä voi olla todennäköistä, että erilaisten analyysien ja mallien hankkiminen ulkoistetaan ja näitä hankitaan niihin erikoistuneilta yrityksiltä. Lisäksi erilaiset analyysit tekevät kaupalliset yritykset saattavat keskittyä tiettyihin aihepiireihin liittyvien suuraineistojen yhdistämiseen ja käsittelemiseen.

Tarkastelluissa tutkimuksissa suuraineistoja käytettiin lähinnä ilmastonmuutoksen vaikutusten arviointiin yritysten liiketoimintaan liittyen. Analyysillä pyrittiin löytämään optimaalisia keinoja sille, miten eri yritysten liiketoiminta voisi jatkua vaikka ilmastonmuutos etenisi. Tutkimukset keskittyivät pääosin ilmastonmuutokseen sopeutumiseen, ei juurikaan ilmastonmuutoksen vaikutusten ennaltaehkäisevään vähentämiseen.

Erilaisia suuraineistoja on käytetty aiemminkin tutkimuksissa ilman että asiaa on erityisesti korostettu (esim. Mäkinen 2019). Tarkastelluissa vuoden 2019 tutkimuksissa suuraineisto -termin (big data) käyttäminen kuvasi ehkä enemmän suuraineistoihin liittyvää yleistä kiinnostusta kuin sitä, että jonkinlainen erityinen teknologiamurros olisi käynnissä. Erilaisia aineistoja kerätään ja muutetaan sähköiseen muotoon yhä enempi.

Todennäköisesti yksi tulevaisuuden kehityssuunta olisi suuraineistojen käsittelyn yhä suurempi automatisointi sekä analysoinnin pilkkominen ja hajauttaminen. On myös todennäköistä, että erilaisia suuraineistoja tulee olemaan yhä useampia, ja näiden

hallinta ja käyttäminen voivat muodostaa ongelmia liittyen sekä laskentatehoon sekä ihmisten kykyyn ymmärtää suuria tietomääriä. Lisäksi ongelmana tulevaisuudessa saattaa olla se, että monimutkaisista analyyseistä voi olla hankala saada sellaista tietoa, jonka tavallinen ihminen ymmärtäisi vaivatta ja tällöin ilmastonmuutoksen hidastamiseksi tehtävät toimet saattavat kärsiä. Datan määrien mahdollinen lisääntyminen voi tarkoittaa myös sitä, että tulevaisuudessa erilaiset analyysit ovat lähinnä vain suurempien liikeyritysten käytettävissä johtuen analyysien vaatimista resursseista (mikäli analyysejä halutaan tehdä itse eikä ulkoistaa).

On myös huomattava, että ilmastonmuutoksen muutokset tapahtuvat suhteellisen pitkän aikavälin kuluessa kun taas yritysten liiketoimintaa saatetaan suunnitella lyhyen aikavälin tavoitteita silmällä pitäen. Tästä voi seurata tulevaisuudessa se, että vain osa yrityksistä alkaa optimoimaan toimintojaan erilaisten suuraineistoanalyysien pohjalta. Tämänäyttelyisiä yrityksiä voisi olla esim. energia-alan yritykset ja erilaiset maatalouteen liittyvät yritykset, joiden liiketoiminta on pitkäjänteistä ja joiden liiketoiminnassa erilaisista pitkän ajan ilmastonmuutosanalyyseistä on hyötyä. Voi olla todennäköistä, että yritykset, joiden liiketoimintaan ilmastonmuutos ei suoranaisesti vaikuta, eivät välttämättä tule tämänäyttelyisiä analyysejä käyttämään.

Käsiteltyjen tutkimusraporttien valossa näyttäisi kuitenkin siltä, että hyvin monimutkaisten ilmiöiden (kuten ilmastonmuutos) ymmärtämiseen ei oikein ole muuta keinoa kuin erilaisten suuraineistojen yhdistäminen ja näiden analysointi sekä ymmärrettävien johtopäätösten tekeminen.

Muutamassa tutkimuksessa (esim. Boas et al. 2019) oli käytetty suuraineistoanalyysin rinnalla myös muita tiedonhankintamenetelmiä (esim. haastattelut) joilla pyrittiin ymmärtämään tutkittua asiaa paremmin sekä saamaan uutta tietoa. Vastaavasti toisessa ääripäässä oli tutkimus, jonka painopiste oli lähinnä pelkästään suuraineiston analysoinnissa ja laskennan tarkkuuden optimoinnissa (Tiampo et al. 2019). Muut tutkimukset olivat näiden kahden väliltä. Mikäli ilmastonmuutoksesta halutaan saada enemmän tietoa, olisi huomioita se, ettei pelkästään suuraineistoja analysoimalla voida saada kaikkea tietoa vaan tiedonhankinta vaatii todennäköisesti useampien suuraineistojen käyttöä sekä usein myös asiantuntijoiden tai muiden asiaan liittyvien henkilöiden haastatteluja. Tällöin tutkittavasta ilmiöstä voidaan saada kattavampi kuva kuin mikäli pelkästään suuraineistojen analysointia olisi käytetty.

Vaikka referoiduissa tutkimuksissa yhdisteltiin useampia suuraineistoja, niin tutkimusten painopiste oli yleensä melko kapealla alueella. Sellaiset tutkimukset, missä olisi yhdistetty useita hyvin laajoja suuraineistoja ja luotu erilaisia malleja monelle tieteenalalle, olivat hyvin harvassa. Suuraineistojen käsittelyn kanssa ilmastonmuutoksen yksittäiset tutkimukset muistuttavat tältä osin enempi yksittäisten palapelin palojen hahmottamista kuin suuren mittakaavan yleiskuvan luomista. Suuraineisto-termiä käytettiin tutkimuksissa enempi ehkä siihen, että tutkimukset halusivat enempi profiloitua treditermillä.

Alla olevassa kuvassa on hahmoteltu avoimen datan aineistojen käyttämistä menneisyydessä, nykyään ja mahdollisia tulevaisuuden kehityskulkuja. Hahmotelma perustuu mielikuvaan, joka tutkituista (vuonna 2019 julkaistuista sekä aiemmin

julkaistuista näitä tutkimuksia tukevista tutkimuksista) välittyi. Mahdollisia tulevaisuuden kehityspolkuja saattaisi olla seuraavat:

- Suuraineistojen käyttö ilmastomuutokseen liittyen integroituu luontevaksi osaksi yritysten tietojärjestelmiä
- Suuraineistoja ei hyödynnetä yhteiskunnan taantumisen vuoksi
- Suuraineistoja hyödynnetään ilmastomuutokseen liittyen yhä enempi poikkitieteellisissä tutkimusryhmissä

	Menneisyys	Nykyhetki	Tulevaisuus?
Aineisto	Kerättyjä kasvinäytteitä, vihkoon kirjattuja lämpötiloja	Avoimia lähteitä sekä muita lähteitä	Erilaisista lähteistä automaattisesti kerätyt ja koostetut tiedot
Tekijät	Kapean alan erikoisosaja	IT-taitoinen erikoisosaja, harrastelijoita yms.	Tekoäly ja ihmiset keräävät ja analysoivat tietoa reaaliajassa
Kattavuus	Data suppealta alueelta	Globaalia ja paikallista dataa	Globaali data, standardirajapinnat yms.
Muut lähteet	Pelkkä data-analyysi	Data-analyysi sekä esim. haastatteluja	Data-analyysijä, haastatteluita, muita poikkitieteellisiä lähteitä
Loppukäyttäjä	Tutkimusorganisaatiot	Tutkimus ja yritykset sekä harrastelijat	Tutkimus, yritykset, harrastelija sekä tekoäly
Kaupallisuus	Vähäinen kaupallinen arvo	Riskienhallinta, kaupunkisuunnittelu yms.	Osa-alueena kaikessa liiketoiminnassa
Kiinnostavuus	Kiinnostaa lähinnä yliopistoja	Ilmastomuutos kiinnostaa yleisesti kaikkia	Elintärkeä huomioitava osa-alue
Toteutus	Erillisiä tutkimuksia eri alojen julkaisuissa	Tutkimuksia, uutisia, verkkosivuja	Interaktiivisia tutkimuksia, uutisia, joukkoistamista

Taulukko 7. Mahdollisia suuraineistojen käytön kehityssuuntia ilmastomuutokseen liittyvässä tutkimuksessa.

Google Earth Engine on vapaasti käytettävä suuraineistoja hyödyntävä työkalu. Mikäli erilaista dataa ja analyysityökaluja on kattavasti kenen tahansa saatavilla tulevaisuudessa, voi tämä (parantunut tiedonsaanti) nopeuttaa huomattavasti suuraineistojen käyttämistä ilmastomuutokseen liittyen. Yksi tutkimus oli käyttänyt Google Earth Engineä (Bastin et al. 2019).

Tässä lopputyössä tarkasteltiin myös erityisesti Ylen uutisointia liittyen referoituihin tutkimuksiin. Tarkastelun perusteella voidaan todeta, että referoiduissa tutkimuksissa esiintyvät aiheet esiintyivät myös Ylen uutisissa. Aiheista oli uutisoitu jo ennen vuotta 2019 kun samantyyllisiä (kuin vuonna 2019 julkaistuja) tutkimuksia oli tehty. Ylen uutisointi keskittyi lähinnä aluespesifisiin ilmastomuutosta käsittelevien tutkimusten aihepiireihin. Alla olevassa taulukossa on hahmoteltu hyvin karkeasti sitä, minkä tyyppisiä uutisia liittyi referoituihin tutkimuksiin. Ylen uutiset on valittu hyvin karkealla otannalla ja lähinnä sen perusteella, että uutisissa on käsitelty samoja aihepiirejä tai

huolia kuin mitä referoidussa tutkimuksessa oli käsitelty.

Referoitu tutkimus	Vastaavantyyppinen Ylen uutisointi
Tarvitaan suuraineistoja, datan ymmärtämistä ja ennustamista (Hassani et al. 2019)	
Earth Observation Data -termi, datan ymmärtäminen keskeistä (Sudmanns et al. 2019)	
Big Earth Data -termi, systeemiarkkitehtuurimalli (Yang et al. 2019)	
Suuraineistojen ja paikallisten ihmisten haastattelujen yhdistäminen (Boas et al. 2019)	
Suuraineistojen käyttäjäryhmät ja käytön motiivit (Lassinantti et al. 2019)	
Suuraineistojen analysoinnilla lisäarvoa kestävä kehityksen firmoille (Singh et al. 2019)	Yritysten viherpesu (Yle 2019a), uusiutuvan energian kannattavuus (Yle 2019b), yritysten ilmastomuutostietoisuus (Yle 2019c)
Kaupungeista kerättyjen suuraineistojen hyödyntäminen (Creutzig et al. 2019)	Yksityisautoilun päästöt ja kaupunkirakenne (Yle 2015a)
Smart Cities -termi, kaupungeista kerättyjen suuraineistojen mahdollisuudet (Allama et al. 2019)	Kiinan valvontayhteiskunta (Yle 2018a), Nesteen biopolttoine (Yle 2019d), ilmastoahdistus (Yle 2019e, Yle 2019f), Espoon ilmastovaikutus (Yle 2018e)
Biodiversiteettidatan erikoispiirteitä (König et al. 2019)	Sukupuuttouhat (Yle 2019g)
Koneoppiminen ja data-analyysien haasteet (Reichstein et al. 2019)	
Suuraineistojen käyttäminen energiateollisuudessa (Li et al. 2019)	
Globaalien datavisualisointien muuttaminen paikalliseksi toiminnaksi (Schneider et al. 2019)	
Äärisäiden ennustaminen (Brönnimann et al. 2019, Muszynski et al. 2019).	Ilmastomuutos yritysten strategioissa (Yle 2019h), tulvien ennakointi (Yle 2018b), metsälampien ilmastotieto (Yle 2015b), arktisen sulamisen vaikutus (Yle 2018c)
Metsäpalojen esiintymistodennäköisyydet Italian eri osissa (Michetti et al. 2019).	Ruotsin metsäpalot (Yle 2019i)
Ilmastomuutoksen riskit Himalajan alueen maanviljelykselle (Shukla et al. 2019).	Ilmastomuutoksen vaikutus Amerikan keskilämmen maanviljelykselle (Yle 2019j)
Maanjäristysten esiintymistodennäköisyydet Kanadan Montrealissa (Tiampo et al. 2019).	Sulawesin vuoden 2018 maanjäristys (Yle 2018d)
Wall Street Journal lehden ilmastomuutosaiheisen sana-analyysin käyttö sijoittamisessa (Engle et al. 2019).	Tekoäly lyhytkestoisessa pörssikaupassa (Yle 2017b)
Ilmastomuutoksen vaikutus turvapaikanhakijoiden määrään (Abel et al. 2019).	Ilmastopakolaiset (Yle 2009).
Arktisen alueen jääpeitteen sulaminen (Bliss et al.	

Referoitu tutkimus	Vastaavantyyppinen Ylen uutisointi
2019).	
Ilmaston lämpenemisen vaikutus jokisuistossa Yhdysvalloissa (Bamber et al. 2019, Chen et al. 2019).	Merenpinnan nousu jäätiköiden sulaessa (Yle 2019k), äärisäiden lisääntyminen (Yle 2011), tulvavaara-alueet Suomessa (Yle 2018f)
Yhdysvaltojen ilmasto-olosuhteiden luokittelu (Sathiaraj et al. 2019).	Ilmastonmuutosvisualisointi (Yle 2018g), autoiluja ja lihansyöntiä vähennettävä Suomessa (Yle 2019m)
Saksan metsien kasvuolosuhteet (Dietrich et al. 2019).	Äärisäiden lisääntyminen (Yle 2018b), siitepöly ja allergiat (Yle 2008)
Brasilian aurinkosäteilyn jakautuminen alueittain tulevaisuudessa (de Jong et al. 2019).	
Suuraineistojen käyttäminen energia-alalla (Li et al. 2019).	
Pohjois- ja Keski-Euroopan sienilajien elinolosuhteet (Andrew et al. 2019).	Ilmastonmuutos sekoittaa ekosysteemejä (Yle 2019n)
Ilmastonmuutoksen vaikutus taloudelliseen menestykseen (Diffenbaugh et al. 2019).	
Ruokahuollon suunnittelu (Srinivasana et al. 2019).	Ilmastonmuutoksen haitat hyötyjä suuremmat Suomessa (Yle 2018h)
Ilmastonmuutoksen vaikutus Los Angelesin sähkötalouteen (Burillo et al. 2019).	Ilmastonmuutos lisää energiankäyttöä (Yle 2019o)
Ilmastonmuutoksen vaikutus karhujen elinolosuhteisiin Espanjassa (Penteriani et al. 2019).	Kasvilajien sukupuutto (Yle 2019p), vieraslajien leviäminen Suomessa (Yle 2019q),
Asiantuntijoiden käyttäminen suuraineistojen rinnalla (Reside et al. 2019).	
Satelliittikuvien käyttäminen puiden kasvatuspaikkojen etsintään (Bastin et al. 2019).	Puuston lisääminen tehokkain torjuntakeino (Yle 2019r),
Useiden suuraineistojen yhdistäminen ilmaston trendien selvittämiseksi (Neukom et al. 2019).	esim. Yhdysvaltojen helleaalto (Yle 2019s).

Taulukko 8. Referoidut tutkimukset (data, riskienhallinta, aluespesifiset) ja vastaavia Ylen uutisia.

Tutkimuksissa ja Ylen uutisissa oli havaittavissa seuraavanlaisia trendejä liittyen ilmastonmuutoksen tutkimukseen suuraineistoja käyttämällä:

- Äärisäiden lisääntyminen ja näihin varautuminen
- Napa-alueiden jääpeitteiden sulaminen ja tähän liittyvät uhkakuvat
- Ilmastonmuutoksen vaikutus maanviljelykseen
- Ilmastonmuutoksen vaikutus biodiversiteetteihin ja eliöiden sukupuuttoon
- Kaupunkien varautuminen ilmastonmuutoksen vaikutuksiin
- Energiantuotannon ja energiahuollon varautuminen ilmastonmuutokseen
- Maanjäristysten ennakointi
- Suuraineistojen määrän kasvu ja hyödyntäminen
- Metsäpalot ja lisääntyvän kuumuuden vaikutus
- Globaalien ilmastomallien ja ennusteiden soveltaminen paikallisesti

- Suuraineistojen käyttäminen muiden tutkimustapojen (esim. haastattelut) rinnalla

Loppupäätelmänä liittyen tutkimuskysymykseen voitaneen todeta, että suuraineistojen käyttäminen on enempi looginen kehityssuunta ilmastonmuutokseen liittyvässä tutkimuksessa. Erilaisia laajoja aineistoja on käytetty aiemminkin tutkittaessa erilaisia pitkän aikavälin muutoksia. Suuraineistot eivät itsessään pelastane ihmiskuntaa ilmastonmuutoksella, mutta erilaisia suuraineistoja käyttämällä ja yhdistelemällä voidaan tähän liittyvää tutkimusta tehdä tehokkaammin. Analysoinnin tuloksia pitäisi myös pystyä hyödyntämään yhteiskunnan päätöksenteossa ja näitä hyödyntämistoimia pitäisi pystyä jotenkin mittaamaan. Koska yhteiskunta on hyvin monimutkainen, tällainen suuraineistojen tulosten hyödyntäminen voi olla hyvin haasteellista ja hyödyntäminen saattaa rajoittua lähinnä erilaisten uhkien ennakointiin, ei niinkään ilmastonmuutoksen systemaattiseen hillitsemiseen kaikin tavoin. Lisäksi yhteiskunnassa voi olla erilaisia intressiryhmiä (esim. suuryritykset), joilla voi olla erilaisia kaupallisia tavoitteita liittyen ilmastonmuutoksen aiheuttamiin erilaisiin ilmiöihin. Erityisesti avoimesti saatavat lähteet mahdollistavat yhä useampien erilaisten ihmisryhmien tekemät analyysit. Haasteena suuraineistojen käyttämisessä on kasvanut vaikeus datamassojen yhdistämisessä ja hallitsemisessa. Lisäksi eri lähteistä saatujen suuraineistojen yhdistäminen vaatii yhä enemmän matemaattisia ja tietoteknisiä taitoja.

Mahdollisia jatkotutkimuksen aihepiirejä olisi esim.:

- Ilmastonmuutoksen ja suuraineistojen hyödyntämisen seuraamisen jatkaminen ja mahdollisten tutkimustrendien syvällisempi analysointi. Tutkimustrendien syvällisempi analysointi tarkoittaisi myös aiempina vuosina tehtyjen tutkimusten analysointia ja tutkimusten määrien analysointia eri aihepiireistä. Tällöin voitaisiin saada selville se, minkä tyyppisten aiheiden määrä on lisääntynyt ja minkä aihepiirien aiheet ovat vähentymässä.
- Yrity maailman käyttämien suuraineistojen hyödyntämisen tutkiminen. Tässä diplomityössä käsiteltiin lähinnä julkisia tutkimuksia, joissa oli tyypillisesti käytetty erilaisia avoimen lähteiden aineistoja. Yrityksiä haastatteleamalla voitaisiin saada tietoa myös erilaisista liiketoimintaan liittyvistä kaupallisista sovelluksista sekä siitä, missä laajuudessa erilaisia suuraineistoja käytetään yrityksissä ilmastonmuutokseen liittyen.
- Suuraineistoihin perustuvien tutkimusten tulosten vertaaminen muilla tavoin (esim. haastatteleamalla) tehtyihin tutkimuksiin. Tällöin voitaisiin saada parempi käsitys siitä, millaisissa tapauksissa suuraineistojen analysointi olisi tehokkain tapa.
- Kattavampi analyysi uutisten ja tutkimusten aiheista, jonka avulla olisi mahdollista selvittää, kuinka tarkasti uutiset seuraavat ilmastonmuutokseen ja suuraineistojen hyödyntämiseen liittyvää tutkimusta.
- Suuraineistoanalyysien tulosten käyttäminen ilmastonmuutoksen liittyvissä hankkeissa yhteiskunnassa.
- Kattavamman otoksen käyttäminen. Tässä diplomityössä keskityttiin alkuvuonna 2019 julkaistuihin tutkimuksiin ja laajempi aineisto voisi lisätä lopputulosten uskottavuutta sekä tuoda tietoa erilaisista pitkän aikavälin trendeistä.

On myös huomioitava, että erilaisten suuraineistojen ja analyysien määrän kasvaminen on vain yksi tulevaisuusskenaario. On myös mahdollista, että ilmastonmuutos muuttaa yhteiskuntaamme tuntemattomalla tavalla tai teknologinen kehitys mahdollistaa suuraineistojen käytön ihan uudella tavalla kuin millä tavalla referoidut tutkimukset olivat käyttäneet suuraineistoja.

5.4 Tutkimuskysymysten vastaukset

Referoidut tutkimukset antoivat läpileikkauksen vuoden 2019 alkupuolella tehdyistä tarkastelluista tutkimuksista. On huomattava, että vain 34 tutkimusta referoitiin, mutta ilmastonmuutoksesta on tehty suuri määrä tutkimuksia. Referoinnin perusteella referoiduissa tutkimuksissa käytetty "big data" -termi viittaisi enempi trendisanojen käyttämiseen kuin suuraineistojen jonkinlaiseen erityisen ihmeelliseen hyödyntämiseen.

Ilmastonmuutos on hyvin monimutkainen ilmiö, jonka tutkimisessa suuraineistoilla on hyvin todennäköisesti tulevaisuudessa suuri rooli. Referoitujen tutkimusten perusteella erityisen kiinnostavia tutkimuksia olivat sellaiset tutkimukset, joissa oli yhdistelty erilaisia suur- ja pienaineistoja sekä erilaisia asiantuntija-arvioita sekä paikallisten ihmisten haastatteluja. Kenties tämäntyyppinen suuraineistojen käyttämistapa on yksi mahdollinen tulevaisuuden kehityssuunta.

Referoitujen tutkimusten määrä oli erittäin pieni verrattuna kaikkeen ilmastonmuutoksesta tehtyyn tutkimukseen. Laajemmasta aineistosta olisi mahdollisesti saatavissa lisää tietoa esim. erilaisista suuraineistojen käyttötavoista liittyen ilmastonmuutoksen tutkimukseen. Ilmastonmuutokseen liittyvästä tutkimuksesta on tehty kirjallisuustutkimuksia jo aiemmin (esim. Haunschild et al. 2016), ja tässä diplomityössä haluttiin keskittyä tarkastelemaan tarkemmin suuraineistojen käyttämistä vuoden 2019 alussa julkaistuissa tutkimuksissa.

- TUTKIMUSKYSYMYS: Miten erilaisia suuraineistoja on käytetty tässä lopputyössä tarkastelluissa ilmastonmuutokseen liittyvässä tutkimuksessa vuonna 2019 julkaistuissa tutkimuksissa?

Tyypillisesti tutkimuksissa on käytetty yhtä tai useampaan (usein julkisesti saatavilla olevaa) suuraineistoa, jota on analysoitu ja tyypillisesti on tehty erilaisia matemaattisia malleja, joilla erilaisia tulevaisuuden kehityskulkuja on arvioitu. Tutkimuksen tilaaja on esim. valtiollinen taho, kaupallinen toimija tai yhteisö. Tutkimukseen liittyy tyypillisesti eri alojen ihmisiä (esim. data-analytikkoja, luonnontieteiden asiantuntijoita yms.). Erilaisia suuria aineistoja on käytetty aiemminkin mutta nykyään aineistojen käyttäminen on helpottunut koska aineistoja on tarjolla yhä enemmän. Käsitellyt tutkimukset kostuivat tyypillisesti itse suuraineistojen käyttämiseen (data-analytiikka ilmastonmuutokseen liittyvässä tutkimuksessa) tai jonkun maantieteellisen alueen ja/tai asian tutkimiseen erilaisia suuraineistoja analysoimalla.

Referoitavissa tutkimuksissa oli tyypillisesti tietynlainen rakenne, joka koostui datan keräämisestä, analysoimisesta, mallien luomisesta, ennustamisesta sekä ennusteiden

analysoinnista (kuvattu esim. Hassani et al. 2019). Tyypillisesti data-aineistosta yritettiin löytää yhteyksiä tiettyjen tutkittavien asioiden välillä, esim. metsäpalot, Italian metsäalueet, ilmaston lämpeneminen alueella sekä alueen sosio-ekonominen tilanne kuten tutkimuksessa Michetti et al. 2019. Suuraineistojen käyttämiseen liittyi myös erilaisia teknisiä toimenpiteitä liittyen tietojen tallentamiseen, tietojärjestelmien käyttämiseen, erilaisten algoritmien valitsemiseen sekä erityisesti visualisointeihin (esim. Yang et al. 2019). Tyypillisessä suuraineiston analysoinnissa saatiin tietoa esim. tietyn alueen tulevaisuudennäkymistä, jolloin saadut tulokset (ikään kuin suuraineistosta analyysin perusteella saatu suuraineisto) visualisoitiin. Hyvä esimerkki tästä oli esim. puiden kasvupaikkoja käsittelevä tutkimus, jossa lopputuloksen visualisoitiin sopivat kasvupaikat maailmankartalle (Bastin et al. 2019). Tyypillisesti referoiduissa tutkimuksissa tekijät olivat tietyn alan tutkijoita, mutta data-analyysien tekemisessä tarvitaan tyypillisesti myös tietotekniikan ammattilaisia (esim. Lassinantti et al. 2019). Koska suuraineistoja on vapaasti saatavilla ja tarvittavia työkaluja, kuten Google Earth Engine, on vapaasti käytettävänä verkossa, periaatteessa kuka tahansa asioihin tarpeeksi perehtynyt voisi tehdä vastaavia analyysijä.

Suuraineistoja käyttämällä saadaan täsmällistä tietoa tyypillisesti tietyltä rajatulta (maantieteellisesti ja aihepiirillisesti) alueelta ja suuraineistojen käyttäminen on enempi palapelin palojen luomista. Vastaavasti palapelin kokoamisessa ja asioiden yhdistämisessä ja ymmärtämisessä tarvitaan todennäköisesti muuntyylisiä tutkimuksia tukemaan suuraineistoilla tehtyjä tutkimuksia.

Joissain tutkimuksissa oli käytetty data-analyysin rinnalla myös muita lähteitä kuten esim. paikallisten ihmisten haastattelemista (Boas et al. 2019). Tällöin on tyypillisesti saatu lisää tietoa tutkittavasta asiasta (kuin pelkkää data-analyysia käyttämällä). Yksi mahdollinen tämäntyylinen kehityssuunta voisi data-analyysien ja erilaisten muiden tutkimustapojen (kuten haastattelut, asiantuntija-arviot yms.) yhdistäminen. Hyvä esimerkki erilaisten lähteiden yhdistämisestä oli tutkimus, jossa yhdistettiin eri tavoin (esim. säätiedot, puiden vuosirenkaista saadut tiedot yms.) kerättyä tietoa ja tätä analysoitiin (Neukom et al. 2019). Erityisesti tämän tyyppinen, monia erilaisia datalähteitä yhdistelevä tutkimus vaikutti hyvin vakuuttavalta ja ehkä tämä voi olla yksi kasvava tulevaisuuden trendi.

Tyypillinen, tässä lopputyössä referoitu, vuonna 2019 julkaistu ilmastomuutosta suuraineistojen avulla tutkiva tutkimus on tyypillisesti maantieteellisesti ja tutkimuskohtaisesti rajattu tutkimus, jossa on analysoitu yhtä tai useampaa (tyypillisesti avoimesti saatavilla olevaa) suuraineistoa, josta on tehty ennusteita tulevalle kehitykselle. Tyypillinen tämänkaltaisen tutkimus muistuttaa palapelin palaa suuremmassa kokonaisuudessa.

Suuraineistoja analysoimalla voitiin löytää erilaisia vuorovaikutussuhteita sekä luoda erilaisia matemaattisia malleja kuvaamaan mahdollisia tulevaisuuden kehityskulkuja. Pienempien aineistojen ja perinteisellä tavalla tehtyjen tutkimusten (esim. yhden tutkijan keräämä suppea aineisto) avulla voi olla hankala päästä kiinni pitkän aikavälin trendeihin ja tarkkojen matemaattisten mallien luominen voi olla hankalaa. Useissa referoiduissa tutkimuksissa lopputuloksena saadut tulokset oli visualisoitu kartalle (esim. karhujen elinolosuhteiden muutos Penteriani et al. 2019). Ehkä yksi keino

tutkimustulosten vaikuttavuuden lisäämiseksi voisi olla interaktiivisuuden lisääminen, jolloin käyttäjä voisi itse säätää jotain yksinkertaisia parametreja sekä havainnoiden näiden vaikutusta tutkittavaan asiaan.

Referoidut tutkimukset olivat tyypillisesti luonnontieteiden alalta, mutta mukana oli myös tutkimus, jossa tarkasteltiin turvapaikanhakijoihin liittyviä asioita suuraineistoanalyysin kautta (Abel et al. 2019) sekä yhdessä tutkimuksessa analysoitiin sanomalehden uutisia sijoitustoiminnan optimointiin liittyen (Engle et al. 2019). Referoidut ilmastomuutokseen ja suuraineistojen käyttämiseen tutkimukset olivat pääosin aluespesifisiä tai suuraineistojen käyttämiseen liittyviä tutkimuksia. Riskienhallintaan liittyviä tutkimuksia oli vähempi, mutta on huomattava että tässä lopputyössä käytiin läpi ainoastaan 34 tutkimusta. Google Scholar -palvelussa on tyypillisesti yliopistojen ja vastaavien tutkimuslaitosten julkaisuja, kun taas riskienhallintaan liittyvät tutkimukset saattavat olla yritysten itsensä tekemiä ja ei-julkisia.

Referoiduissa aluespesifisissä ja riskienhallintaan liittyvissä tutkimuksissa olleita teemoja (ilmaston lämpeneminen, jäätikköjen sulaminen yms.) käsiteltiin myös Ylen uutisissa. Joistain referoiduista tutkimuksista oli myös lähes samaan aikaan julkaistu uutinen Ylen sivuilla (esim. Bastin et al. 2019 ja Yle 2019r). On huomattava, että myös uutisoinnissa ilmastomuutos on ollut yhä enemmän esillä, joten väistämättä samat aiheet toistuvat jossain muodossa myös uutisissa. Tutkimusten ja Ylen uutisten välittämät ydinviestit (esim. arktinen alue lämpenee tms.) vastasivat hyvin toisiaan.

Referoiduissa tutkimuksissa käytetyt suuraineistot olivat 4V-mallin (IBM 2019) mukaisia suuraineistoja, siltä osin kuin mallia voitiin käyttää suuraineistojen tunnistamisessa. Tutkimusten suuraineistojen sisältämän datan määrä oli suuri, dataan liittyi erilaisia epävarmuuksia, dataa oli monessa referoidussa tutkimuksessa useasta lähteestä ja datassa oli paljon erilaisia muutoksia. Erityisen mielenkiintoiselta vaikutti ilmastotutkimus, jossa oli yhdistetty puiden vuosirenkaista saatua tietoa perinteisempiin säähavaintoihin pitkältä ajalta (Neukom et al. 2019) sekä suuraineistoanalyysit, joissa suuraineistojen rinnalla oli käytetty myös pienaineistoja kuten esim. haastatteluja (Reside et al. 2019, Shukla et al. 2019, Abel et al. 2019, Boas et al. 2019).

- ALIKYSYMYS: Millä tavalla suuraineistot auttavat ilmastomuutokseen liittyvien haasteiden ymmärtämisessä?

Tyypillisesti ilmastomuutoksessa tutkitaan monimutkaisen järjestelmän (esim. ilmaston muuttumisen vaikutusta johonkin tiettyyn asiaan). Yhdessä referoitavassa tutkimuksessa tarkasteltiin ilmastomuutoksen vaikutusta Himalajan alueen maanviljelykseen (Shukla et al. 2019). Tämän tyylisessä tutkimuksessa on oltava paljon tietoa siitä, miten tutkittava asia (alueen ilmasto, jäätiköt, vesistö) on hitaasti muuttumassa. Välttämättä ei ole tiedossa, että minkälaiset takaisinkytkennät ja mallit kuvaavat muutosta, jolloin järkevä lähestymistapa on analysoida kerättyä tietoa (esim. alueen säätiedot ja tiedot maanviljelykseen liittyvistä asioista). Näitä analysoimalla voidaan luoda matemaattisia malleja. Mallien rinnalla voidaan käyttää muita tutkimustapoja (kuten maanviljelijöiden haastattelut Himalajan alueen tutkimuksessa),

jotka voivat tuoda sellaista tietoa, jota esim. pelkkä ilmastonmallien matemaattinen käyttäminen antaisi. Tekemällä tulevaisuuteen liittyviä simulaatioita, voidaan myös luoda erilaisia graafisia visualisointeja, joista nähdään, miten monimutkainen muutos vaikuttaa eri alueilla (esim. missä Himalajalla minkäkin kasvin viljelymahdollisuudet paranevat ja missä vastaavasti viljelyn edellytykset heikentyvät).

- ALIKYSYMYKSET: Minkälaisia keinoja suuraineistot tarjoavat ilmastonmuutokseen liittyvässä tutkimuksessa verrattuna pienaineistoihin?

Tutkittavat ilmiöt ovat erittäin monimutkaisia, eikä pienaineistojen käyttämisellä välttämättä saada tarpeeksi tarkkoja mallinnuksia siitä, miten monimutkaiset järjestelmät toimivat. Lisäksi suuraineistoja analysoimalla voidaan havaita erilaisia säännönmukaisuuksia (patterns) ja vuorovaikutuksia, jotka muuten jäisivät todennäköisesti havaitsematta. Esimerkkinä voidaan käyttää vaikka ilmastonmuutoksen vaikutusten tutkimista Los Angelesin sähkönkulutukseen (Burillo et al. 2019). Tutkittavana on suuri alue, jossa on erilaisia rakennuksia ja mahdollisesti erilaisia sähkökäyttäjryhmiä. Lisäksi tarvitaan jonkinlainen malli siitä, miten ilmasto muuttuu ja millä tavoin sähkönkulutus riippuu lämpötiloista. Lisäksi tutkimuksen tuloksena ei haluta yhtä tiettyä arvoa tai johtopäätöstä, vaan lopputulokset ovat todennäköisesti rakennus- ja sähkökäyttäjryhmäkohtaisia. Jos samaa tutkittaisiin pienemmillä aineistoilla esim. haastatteleamalla ihmisiä, niin ihmiset eivät välttämättä itse edes tiedosta sähkökäyttötapoja eivätkä osaa ennakoita sitä, miten ilmasto heidän alueellaan muuttuu. Lisäksi erilaisia kehityspolkuja varten tarvitaan kuitenkin jonkinlainen ennuste tulevaisuudesta, ja esim. ilmastomalleja luotaessa tarvitaan käytännössä aina yksi tai useampi suuraineisto, joita analysoimalla voidaan luoda ennustemalleja tulevaisuudelle.

- ALIKYSYMYKSET: Miksi suuraineistoja käytetään ilmastonmuutokseen liittyvässä tutkimuksessa?

Ilmastonmuutokseen liittyvä tutkimus on kasvanut erityisesti 1990-luvulla (Haunschild et al. 2016), jolloin myös tietotekniikan ja tietoverkkojen käyttäminen yleistyi laajemmassa mittakaavassa. Aiemmin erilaisia mittaustietoja oli ehkä saatavilla vain rajoitetulle yleisölle ja tarvittava osaaminen sekä tietokoneiden laskuteho asetti todennäköisesti rajoituksia. Osaamisen ja teknologian kehittymisen sekä ilmastonmuutokseen liittyvän kiinnostuksen kasvaessa myös ilmastonmuutokseen liittyvä tutkimus on kasvanut. Lisäksi erilaiset vapaasti käytettävät palvelut (kuten Google Earth Engine) ovat mahdollistaneet sen, että kuka tahansa voi analysoida erilaisia suuraineistoja ja tehdä erilaisia mallinnuksia.

Ilmastonmuutoksessa tarkastellaan hyvin monimutkaisia ilmiöitä, joista ei välttämättä ole mahdollista saada tietoa muuten, kuin analysoimalla suuria datamääriä ja luomalla niistä erilaisia matemaattisia malleja, joilla erilaisia tulevaisuuden kehityspolkuja voidaan ennustaa. Esimerkkinä on vaikka etsiminen puille vapaita kasvupaikkoja (kuten Bastin et al. 2019). Etsinnässä tarvitaan tietoa maapallon eri alueiden kasvuolosuhteista nyt ja tulevaisuudessa sekä muita vastaavia tietoja. Lopputuloksena on myös suuri määrä dataa, joka voidaan visualisoida (esim. näytetään sopivat kasvupaikat kartalla,

josta mahdollisesti asiantuntijat, poliitikot, paikalliset ihmiset yms. voivat tehdä omia päätöksiä).

- ALIKYSYMYS: Näkyykö Suomen Ylen vuonna 2019 julkaistuissa uutisissa vastaavia teemoja ja asioita kuin käsitellyissä tutkimuksissa?

Ilmastonmuutos on ollut alkuvuonna 2019 voimakkaasti uutisotsikoissa ja uutisissa. Erilaisia ilmastoon ja luontoon liittyviä ilmiöitä ja muutoksia on raportoitu uutisissa viikoittain. Uutisten mukaan yhä useampi yritys on alkanut reagoimaan ilmastonmuutokseen ja ilmastonmuutoksesta on tullut osa yhä useamman yrityksen strategiaa. Kuitenkin käsitellyissä tutkimuksissa kaupallisuus ja liiketoiminta näkyi lähinnä siten, että tutkimuksissa pohdittiin lähinnä sitä, miten liiketoiminta saataisiin jatkumaan ilmastonmuutoksesta huolimatta.

Ylen uutisten aiheissa näkyi samoja aiheita kuin mitä oli referoiduissa tutkimuksissakin. Vastaavantyyllisiä uutisia oli julkaistu myös aiemmin samankaltaisista tutkimuksista, kuin mitä tutkimuksia vuonna 2019 oli julkaistu. Ylen uutiset käsitelivät erityisesti aluespesifiset -kategorian tutkimuksia. Referoituja tutkimuksia ja Ylen uutisten sisältöjä vertaamalla voidaan päätellä, että Ylen uutiset ovat luotettavia ja totuudenmukaisia tältä osin.

- ALIKYSYMYS: Pelastavatko suuraineistot meidät ilmastonmuutokselta – mikä on suuraineistojen rooli nykyisessä ilmastonmuutokseen liittyvässä tutkimuksessa?

Pelkästään suuraineistoihin perustuvat tutkimus tuskin riittää ilmastonmuutoksen ymmärtämisessä tai ilmastonmuutoksen hidastamisessa. Vaikka tutkimuksissa yleinen trendi on erilaisten data-analytiikan hyödyntäminen ja aineistojen datamäärien kasvaminen, niin pelkästään suuraineistojen analysoinnilla ei saatane yleiskuvaa. Tilannetta voisi verrata auton korjaamiseen, jossa tunnetaan joku satunnainen määrä autonomisia ja niiden toiminta, mutta itse autolla ajaminen ja liikennesääntöjen tunteminen voi olla osaltaan arvoitus. Tutkimuksissa korostui enempi pienien osa-alueiden (auton osat) tunteminen mutta vastaavissa Ylen uutisissa käsiteltiin näiden lisäksi myös suurempia kysymyksiä (liikennesäännöt).

Lisäksi pelkästään suuraineistoja käyttämällä (ilman ymmärrystä taustalla vaikuttavista asioista) osa tutkittavasta tiedosta jää todennäköisesti löytymättä. Suuraineistojen analysoimiseen on hyödyllistä käyttää esim. erilaisten asiantuntijoiden haastatteluja jotta tutkittava ilmiö ja data ymmärretään oikein. Suuraineistot ovat enempi yksi täydentävä osa tutkimusta.

Suuraineistojen avulla voidaan kuitenkin hahmottaa suurista datamassoista erilaisia riippuvuuksia ja vuorovaikutussuhteita. Referoiduissa tutkimuksissa oli pystytty analysoimaan erilaisia, tyypillisesti luonnonilmiöistä kerättyä mittausdataa. Kuitenkin, koska tarkasteltavat ilmiöt olivat hyvin monimutkaisia, suuraineistojen analysointiin liittyy tyypillisesti myös erilaisia epävarmuuksia. Välttämättä ei tiedetä, mitkä

mitattavat asiat liittyvät ilmaston käyttäytymiseen ja välttämättä ei tiedetä myöskään sitä, sisältääkö käytetty data tutkimuksen kannalta oleellista tietoa. Lisäksi välttämättä ei ole edes käsitystä siitä, mitä tietoa puuttuu ja kuinka luotettavaa ja kattavaa tieto on. Vaikka ilmasto muuttuu pitkän aikavälin kuluessa hitaasti, näitä hitaita muutoksia voi olla myös hankala havaita lyhytaikaisten muutosten joukosta. Mikäli käytössä on myös kattavaa ymmärrystä ja näkemystä siitä, millaisia asioita suuraineistosta on tarkoitus etsiä ja millaista tietoa suuraineisto mahdollisesti sisältää, voidaan suuraineistoa analysoimalla saada uutta tietoa ja erilaisia simulaatiomalleja.

Referoiduissa tutkimuksissa painopiste oli ehkä enempi erilaisten ilmiöiden (esim. jäätiköiden sulaminen) havainnointi ja ennustaminen, kuin se, että miten erilaisia ilmaston lämpenemiseen liittyviä ilmiötä voidaan ennaltaehkäistä. Voi olla, että ennaltaehkäisy ja ennakoointi vaatii lisää tutkimusta ja ilmastoon liittyvien asioiden ymmärtämistä. Lisäksi ilmaston lämpenemiseen reagoimisessa voi ongelmana olla se, mitä reagoinnista seuraa. Ylen uutisten mukaan SSAB on kehittänyt menetelmän, jossa teräksen valmistuksesta tulee päästötöntä mutta samalla sähkönkulutus kasvaa kymmenkertaiseksi (Yle 2019t). Yksi ilmastonmuutokseen liittyvän tutkimuksen painopiste on päästöjen vähentäminen. SSAB:n tapauksessa kymmenkertainen sähkönkulutus voitaisiin kompensoida tuottamalla terästä kymmenen kertaa vähemmän. Tästä seuraisi kuitenkin markkinoilla todennäköisesti puute teräksestä ja puute voisi hidastaa maailmantalouden kasvua. Samalla tavalla erilaiset suuraineistojen analysoinnit voivat todennäköisesti tuoda sellaista tietoa ja sellaisia toimenpide-ehdotuksia, joiden toteuttaminen voi olla hankalaa nykyisessä yhteiskunnassa. Jos suuraineistojen (analysoinnin) halutaan pelastavan ihmiskunnan ilmastonmuutokselta, niin pitäisi ensin varmistaa, että ihmiskunta edes haluaa pelastua (ts. tehdä tarvittavia, oletettavasti melko radikaalejakin toimenpiteitä asian suhteen). Referoitujen tutkimusten perusteella kuitenkin suuraineistojen analysointi on kuitenkin merkittävässä roolissa ilmastonmuutokseen liittyvässä tutkimuksessa ja ratkaisujen etsimisessä.

Yle uutisoi raportista, jonka mukaan ilmastonmuutokseen sopeutuminen maksaisi 1800 miljardia dollaria mutta raha tulisi moninkertaisena takaisin (Yle 2019u, Global Commission On Adaptation 2019). Referoidut tutkimukset antavat suuntalinjoja sille, millaisia vaikutuksia ilmastonmuutoksella olisi maapallolla. Ehkä seuraava askel voisi olla se, miten erilaisia analyysien tuloksia sovelletaan yhteiskunnassa. Yhteiskunnan toiminta voi kuitenkin olla hyvin monimutkaista ja tarvittavat toimenpiteet (esim. kulutuksen vähentäminen) voivat aiheuttaa yhteiskunnassa vastustusta. Ylen uutisissa mainittu 1800 miljardin dollarin käyttäminen voisi tarkoittaa myös varallisuuden uudelleen jakamista ilmastoystävällisemmille yrityksille ja nykyisillä yrityksillä voi olla suuri vaikutusvalta yhteiskunnassa ja mahdollisesti vaikutusvaltaa saatettaisiin haluta käyttää estämään vaurauden siirtyminen. Lisäksi toimenpiteiden toteutus vaatisi myös valtioilta yhteistä tahtoa ja tuloksia pitäisi pystyä jotenkin mittaamaan. Ylen uutisen mukaan Suomella on lähes kaksi sataa kestävän kehityksen tavoitetta mutta vain joka toisen tavoitteen toteutuminen on pystytty seuraamaan (Yle 2019v). Joseph Tainterin mukaan yhteiskuntien kasvava monimutkaistuminen saattaa johtaa siihen, etteivät yhteiskunnat pysty reagoimaan muutoksiin ja monimutkaisen yhteiskuntarakenteen ylläpito vie lopulta liikaa resursseja (Tainter 1998, sivut 118-123). Ehkä Tainterin ajatuksia mukaillen voisi ajatella, että kenties nykyisen länsimaisen yhteiskunnan monimutkaiset rakenteet on tehty nykyistä kehitystä ja talouskasvua ylläpitämään.

Vaikka suuraineistoanalyysien perusteella olisi tietoa tarvittavista toimenpiteistä, ehkä näiden toimenpiteiden toteuttaminen (esim. kulutuksen voimakas vähentäminen) voi olla hankalaa mikäli nykyinen yhteiskunta halutaan pitää samaan aikaan ikään kuin kasassa. Lisäksi voi olla, että tarvittavia muutoksia on hankala tehdä, koska muutokset voivat olla ristiriidassa erilaisten toimijoiden (esim. yritykset, valtaapitävät) intressien kanssa.

Financial Times -lehdessä julkaistiin syyskuussa 2019 poikkeuksellinen kirjoitus, jossa nykyistä talousjärjestelmää kritisoitiin. Kirjoituksessa haluttiin laajentaa yritysten tehtävää voiton tuottamisesta omistajille myös asiakkaiden, työntekijöiden ja yhteisön edusta huolehtimiseen. 180 yhdysvaltalaisista suuryritysten johtajaa (mm. Apple ja Amazon) olivat allekirjoittaneet kannanoton, jossa yritykset sitoutuivat huomioimaan yrityksen taloudellisen tuloksen tekemisen lisäksi työntekijöiden palkat ja koulutukset, ympäristönsuojelun, eettisen toiminnan ja paikallisten yhteisöjen tukemisen (Financial Times 2019, Helsingin Sanomat 2019a, Helsingin Sanomat 2019b). Nykyinen talousjärjestelmä perustuu suurelta osin jatkuvaan talouskasvuun ja ehkä talousjärjestelmän on tulevaisuudessa pakosta oltava kestäväällä pohjalla ja jaettava hyötyä yhä suuremmalla joukolla. Ehkä suuraineistojen rooli on tämentyylisessä kehityksessä antaa yleiskuvaa siitä, mihin suuntaan ilmasto ja ihmiskunta ovat menossa. Kun erilaisten suuraineistojen analysointi tuo yhä enemmän tietoa, tämän tutkimustiedon sivuuttaminen voi olla aina vain vaikeampaa ja kenties jossain vaiheessa yritystenkin pitää päättää, miten ilmastonmuutoksen, maapallon rajallisten resurssien ja muiden vastaavien rajoitukset pitää ottaa huomioon yrityksen toiminnassa.

Ehkä yksi tulkinta suuraineistojen roolista ilmastonmuutoksen hillitsemisessä voisi olla (vapaasti tulkittuna referoitujen tutkimusten ja Ylen uutisoinnin perusteella) esim. seuraavanlainen:

- Erilaisen tiedon ja datan kerääminen suuraineistoiksi. Haasteina voi olla järkevän ja luotettavan datan löytäminen.
- Edellä kuvatun suuraineistojen analysointi maapallon ilmastojärjestelmän ymmärtämiseksi sekä erilaisten riippuvuuksien ja vuorovaikutuksien selvittäminen. Täydentävän tiedon (esim. asiantuntijahaastattelut) käyttäminen erilaisten analyysien rinnalla.
- Erilaisten tulevaisuusskenaarioiden muodostaminen suuraineistoanalyysien perusteella tehdyistä matemaattisista malleista.
- Yhteiskunnan toiminnan ohjaaminen siten, että ilmastonmuutokseen ollaan varautuneita ja ilmastonmuutosta saataisiin hillittyä.
- Tehtyjen toimenpiteiden vaikutusten selvittäminen ja suuraineistojen tietojen täydentäminen.

5.4.1 Kirjallisuustutkimuksen merkitys ja löydöt

Tämän kirjallisuustutkimuksen merkitys on lähinnä antaa yleiskuva siitä, miten suuraineistoja on käytetty vuonna 2019 julkaistuissa tutkimuksissa sekä syventää ilmastonmuutoksen tutkimukseen liittyvää tietoa.

Ilmastonmuutokseen liittyvien tutkimusten lukumääristä on saatavilla helposti tietoa

vaikka Google Scholarin hakutulosten lukumäärien perusteella. Ilmastomuutokseen liittyvästä tutkimuksesta (koskien sekä suur- että myös pienaineistoja käyttäviä tutkimuksia) on tehty jo aiemmin kirjallisuuskatsauksia laajemmalta aikaväliltä. Koska teknologia ja maailmantilanne kehittyy nopeasti, tässä kirjallisuuskatsauksessa haluttiin keskittyä mahdollisimman kohdistetusti nykyiseen tilanteeseen.

Tämän kirjallisuustutkimuksen löytöinä voidaan ehkä pitää sitä, että potentiaalinen kehityssuunta saattaisi olla erilaisten suur- ja pienaineistojen yhdistäminen ja sitä kautta on todennäköisesti mahdollista saada kattavampaa tietoa ilmastomuutoksesta. Lisäksi ilmastomuutosta käsittelevä uutisointi Ylessä tuntuisi vastaavan hyvin sisällöltään sitä, mitä myös vastaavissa tutkimuksissa on käsitelty. Uutisissa ja tutkimuksissa toistuvat samat huolenaiheet.

Tämän kirjallisuustutkimuksen perusteella erilaiset suuraineistoanalyysit näyttäisivät olevan erittäin tehokas työkalu erilaisten ilmastomuutokseen liittyvien tutkimuskysymysten selvittämisessä. Analysoimalla suuraineistoja käyttäviä ilmastomuutostutkimuksia pidemmältä aikaväliltä, olisi analyysin luotettavuutta ja kattavuutta voitu lisätä sekä saada käsitys myös erilaisista trendeistä.

Teknisten työkalujen ja laskentatehojen sijaan ilmastomuutokseen liittyvissä suuraineistoanalyysissä ehkä suuremmat rajoitteet tulee ilmiöiden monimutkaisuudesta. Kun suuraineistoja sekä analyysityökaluja on kenen tahansa saatavilla verkossa, niin ilmiöiden tutkimiseenkin on todennäköisesti käytettävissä monipuolisemmat resurssit kuin aikaisemmin.

5.5 Suuraineistot ilmastomuutoksen jälkeen?

Referoitujen tutkimusten mukaan ilmastomuutos vaikuttaa hyvin moneen asiaan. Ehkä ilmastomuutos vaikuttaa myös suuraineistojen käyttämiseen tulevaisuudessa kun ilmastomuutos on mahdollisesti yltynyt ja tästä syystä tämän diplomityön lopuksi lienee syytä pohtia myös erilaisia suuraineistojen tulevaisuusskenaarioita.

Tässä kappaleessa pohditaan sitä, miten suuraineistoja hyödynnetään tulevaisuudessa ilmastomuutokseen liittyvässä tutkimuksessa. Kappaleen pohdinnat perustuvat löyhästi referoituihin tutkimuksiin sekä tutkimuksissa esiin tulleisiin asioihin.

5.5.1 Suuraineistojen sukupuuttoaalto -skenaario

Referoitujen tutkimusten mukaan ilmastomuutos muuttaa ihan oikeasti kaiken.

Suuraineistoja hyödyntävät ilmastomuutosta tutkivat tutkimukset esittävät yleensä näkemyksen, että tulevaisuudessa datan määrä kasvaa ja erilaiset analyysit monimutkaistuvat ja tarkentuvat. Väite toistuu jossain muodossa käytännössä jokaisessa suuraineistoja käsittelevässä tutkimuksessa tai muussa uutisessa. Väitteelle löytyy myös runsaasti perusteluja, mutta väitettä on harvemmin kiistetty. Tässä kappaleessa on pohdittu luovasti erilaisia kehityskulkuja sille, mitä pitäisi tapahtua jotta erilaisia suuraineistoja ei tulevaisuudessa käytettäisi esim. ilmastomuutokseen liittyvässä tutkimuksessa. Pohdinta on ihan puhdasta mielikuvituksellista spekulatiota, joka

perustuu arvioihin siitä, miten ilmastonmuutos voisi tulevaisuuteen ja suuraineistojen käyttämiseen vaikuttaa.

On kuitenkin mahdollista, että tämä tulevaisuuden kehityssuunta on väärä ja asiat kehittyvät eri tavoin. Erilaisten sensorien (esim. satelliittien kuvauslaitteistot) erottelukyky on parantunut ja nykyään saadaan yhä enemmän ja yhä tarkempaa dataa. Kuitenkin vaikka datan määrä kasvaa (esim. satelliiteista saatavat kuvat ovat yhä tarkempia), niin ehkä jossain vaiheessa sensorien tarkkuuden lisääminen ei enää tuo lisää informaatiota tutkittavista asioista. Lisäksi kasvavien datamäärien tallentaminen ja käsittely vie todennäköisesti yhä enempi energiaa ja aikaa. Mikäli ilmaston lämpeneminen kiihtyy (kuten käsitelty tutkimukset arvelivat), voi olla, että luonnonvarojen kulutusta ja talouskasvua ylipäättään halutaan rajoittaa. Talouskasvun rajoittaminen voisi tarkoittaa suuraineistojen kohdalla sitä, että erilaisten sensoreiden määrä ei kasvaisikaan loputtomiin. Sen sijaan, että sensorit lähettäisivät kaiken keräämänsä raakadatan, kenties tulevaisuudessa erilaiset sensorit harkitsevat tarkkaan mitä dataa kannattaa lähettää ja ehkä erilaisia analyysejä tehdään jo sensoritasolla. Mahdollinen talouskasvun rajoittaminen ilmastonmuutoksen hillitsemiseksi voi tarkoittaa myös sitä, että erilaisille suuraineistoille on yhä vähemmän käyttäjiä. Vastaavasti jos talouskasvua ei rajoiteta ja ilmastonmuutosta ei saada hillittyä, ihmiskunta saattaa keskittyä ilmastonmuutoksen vaikutuksista selviämiseen ja talouskasvun ylläpitämiseen. Jos tulevaisuudessa valtioiden on valittava sääsatelliitin avaruuteen lähettämisen ja työttömyyskorvausten maksamisen välillä, niin ehkä valtiot jättävät osan suuraineistoista keräämättä. Epävarmassa ja eriarvoisessa ilmastonmuutoksen jälkeisessä maailmassa toiminnan painopiste saattaa olla jokapäiväisessä selviämisessä eikä pitkän aikavälin toiminnan suunnittelemisessa suuraineistoja analysoiden. Nykyään länsimaissa esim. matkapuhelin on hyvin yleinen ja erilaisia sosiaalisia medioita käytetään runsaasti. Mahdolliset ilmastonmuutoksen vaikutukset talouskasvuun ja esim. häiriöt energiantuotannossa voisivat johtaa siihen, että tulevaisuudessa sosiaalista mediaa ja verkkopalveluita käytettäisiin vähemmän tai kyseiset palvelut olisivat karsittuja tai niihin olisi pääsy vain pienellä osalla ihmisistä. Todennäköisesti jo kerättyjä suuraineistoja pyrittäisiin silti hyödyntämään.

Erilaiset tutkimukset esittävät tyypillisesti yhteiskunnan teknologisen kehityksen alati jatkuvana ja kasvavana. Kuitenkin ihmiskunnan historiassa on hyvin vähän esimerkkejä siitä, että jokin sivilisaatio (esim. Rooman ikuiseksi tarkoitettu valtakunta) olisi kasvanut jatkuvasti. Tyypillisesti sivilisaatioilla on ollut nousu-, kukoistus- ja romahdusvaihe (mukaillen Diamond 2010). Myös Joseph Tainter on analysoinut eri sivilisaatioiden romahdusta ja esittänyt yhtenä syynä yhteiskunnan lisääntyvän monimutkaisuuden, jonka ylläpitämiseen ei lopulta ole riittänyt resursseja (Tainter 1998, sivut 118-123). Suuraineistojen ja tähän liittyvä teknologia on hyvin monimutkaista ja Tainteria mukaillen voisi olla loogista, että ehkä tulevaisuuden yhteiskunnassa ei kyetä ylläpitämään suuraineistojen hallitsemiseen liittyvää infrastruktuuria kattavasti. Lisäksi Ylen uutisen mukaan maapallon luonnonvarojen ylikulutuspäivä oli vuonna 2019 29. heinäkuuta (Yle 2019w). Jos nykyinen sivilisaatiomme (länsimainen moderni ja globaali yhteiskunta) seuraa menneiden sivilisaatioiden kehityskulkua (idea on esitetty esim. lähteessä Greer 2008), niin ehkä jossain vaiheessa luonnonvarojen rajallisuus, ilmastonmuutos ja muut vastaavat seikat romahduttavat oman sivilisaatiomme. Ilmastonmuutoksesta on uutisoitu monia

vuosia mutta esim. edellä mainittu ylikulutuspäivä tulee joka vuosi yhä aiemmin, koska talouskasvu ja erilaisten tuotteiden valmistus vie uusiutumattomia luonnonvaroja. Tyypillisesti sivilisaatioiden romahdus on kestänyt satoja vuosia ja ehkä romahdusvaiheessa hylätään osa teknologisista innovaatioista ja ehkä siirrytään käyttämään yksinkertaisempia ja paikallisesti saatavilla olevia resursseja (esim. Suomesta löytyy puuta, kiveä ja savea, joista on rakennettu erilaisia perinnerakennettuja taloja vuosituhansien ajan). Digitaalisia suuraineistoja kerätään ja analysoidaan käyttäen monimutkaisia elektronisia laitteita. Ehkä yhteiskunnan romahtaessa ja kenties yksinkertaistuessa erilaisten sensorien määrä vähenee ja myös tarve monimutkaisille suuraineistanalyysille vähenee (ei ole välttämättä enää sellaista liiketoimintaa, joka analyysijä tarvitsisi). Luonnonvarojen säästäminen voisi tarkoittaa kulutuskeskeisestä kulttuurista luopumista ja keskittymistä muihin asioihin. Sosiaalinen media saattaisi korvautua erilaisilla virtuaalitodellisuuteen perustuvilla simulaatioilla, joiden avulla esim. ihmisten liikkumisen tarve työpaikoille ja lomamatkoille vähenisi. Liikkumisen tarpeen radikaali väheneminen tarkoittaisi myös liikenteeseen liittyvän sensoridatan vähenemistä ja liikenteeseen liittyvien monimutkaisten data-analyyysien tarpeen vähenemistä. Nykyisen talousjärjestelmän korvautuminen toisella talousjärjestelmällä voisi vähentää tarvetta erilaisille talouteen perustuville analyyseille. Vastaavasti erilaisten älytalojen sijaan saatettaisiin palata esim. hengittävään perinnerakennustekniikkaan, jossa ei tarvita ylimääräistä elektroniikkaa ja jossa on todennäköisesti huomattavasti vähemmän erilaisia sensoreita. Ehkä ilmastonmuutoksen jälkeiseen tulevaisuuteen selviää osa nykyisestä teknologiasta ja osa käytettävästä teknologiasta perustuu erilaisiin uusiutuviin ja perinteisiin teknologioihin (esimerkkinä vaikka edellä mainitun perinnerakentamisen lisääntyminen paikallisten käytettävien resurssien ja yksinkertaisuuden vuoksi).

5.5.2 Suuraineistojen integroituminen -skenaario

Tässä skenaariossa nykyinen kehitys jatkuu entisellään ja suuraineistojen käyttö ilmastonmuutoksessa noudattelee samoja kehityspolkuja tyypillisten tietoteknisten innovaatioiden kanssa.

Toinen mahdollinen kehityssuunta voisi olla data-analytiikan yleistyminen ja muuttuminen osaksi yritysten ja tutkimuslaitosten toimintaa. Joissain tutkimuksissa (esim. Braun 2015) mainittiin, ettei yrityksillä ole välttämättä vielä standardoitua tapaa hyödyntää ja analysoida suuraineistoja, tai hyödyntää data-aineistoja uuden liiketoiminnan luomisessa (esim. Fred 2018). Mikäli suuraineistojen määrä kasvaa ja analysointityökalut yksinkertaistuvat, todennäköisesti yritykset ja tutkimuslaitokset saattavat tehdä yhä enemmän data-analytiikkaan liittyviä kokeiluja suuraineistojen hyödyntämiseen liittyen. Tällöin suuraineistojen käyttö saattaisi standardisoitua ja käsitys suuraineistojen hyödyntämisestä ilmastonmuutokseen liittyvissä asioissa kehittyä. Tässä kehityssuunnassa suuraineistojen hyödyntäminen voisi olla yhtä luonteva osa esim. yritysten liiketoimintaa siinä missä esim. erilaisten toiminnanohjausjärjestelmien, asiakasrekisterien ja muiden käyttäminen on.

5.5.3 Poikkitieteelliset suuraineistot -skenaario

Ilmastonmuutos on niin poikkeuksellinen ilmiö, että ilmastonmuutoksen aiheuttamia haasteita aletaan ratkaista poikkitieteellisin keinoin.

Joissain referoiduissa tutkimuksissa (esim. Abel et al. 2019) tutkittiin hyvin monimutkaisia ilmastomuutoksen aiheuttamia ilmiöitä kuten esim. ilmastomuutoksen äärisäiden vaikutusta maiden sisäisiin levottomuuksiin. Joissain (esim. Tiampo et al. 2019) taas pääpaino oli data-analyysin teknisessä toteutuksessa matemaatikon näkökulmasta. Yksi kehityssuunta voisi olla suuraineistojen käyttäminen yhä enempi erilaisessa poikkitieteellisessä tutkimuksessa. Aiemmin biologit ovat tutkineet eliöitä, geologit maaperää ja esim. ilmastotutkijat ilmastoa. Koska ilmastomuutos on hyvin monimutkainen ilmiö, tarvitaan todennäköisesti näiden kaikkien tieteenalojen yhteistä näkemystä ja tietoa, jotta käytettävistä suuraineistoista saadaan hyödyllistä tietoa. Joissain tutkimuksissa oli yhdistelty suuraineistojen analysointia paikallisten ihmisten haastattelemiseen (Shukla et al. 2019) tai yhdistelty esim. säädataa ja sosioekonomista dataa (Michetti et al. 2019). Ilmastotutkija ei välttämättä ymmärrä sosioekonomisia asioita, jolloin mahdollinen tulevaisuuden kehityssuunta saattaisi suosia poikkitieteellistä tutkimusta. Tämä (poikkitieteellinen tutkimus) koskisi nimenomaan suuraineistojen käyttämistä monimutkaisten ilmiöiden (kuten ilmastomuutos) tutkimisessa.

Nature Climate Change -julkaisussa on visioitu tämäntyylistä kehityspolkua, jossa suuraineistojen tehokkain hyödyntämistapa olisi sosio-ekonominen suuraineistojen analysointi (Knüsel et al. 2019). Tutkimuksesta ei valitettavasti ollut mahdollista saada kuin tiivistelmä maksumuurista johtuen.

LÄHTEET

Abel Guy J., Brottrager Michael, Cuaresma Jesus Crespo, Muttarak Raya (2019): Climate, conflict and forced migration. *Global Environmental Change*. Volume 54, January 2019, Pages 239-249. Saatavilla (viitattu 16.7.2019): <https://www.sciencedirect.com/science/article/pii/S0959378018301596>

Allama Zaheer, Dhunnyb Zaynah A. (2019). On big data, artificial intelligence and smart cities. *Cities* 89 (2019) 80–91. Saatavilla (viitattu 21.5.2019): <https://www.sciencedirect.com/science/article/pii/S0264275118315968>

Andrew Carrie, Büntgen Ulf, Egli Simon, Senn-Irlet Beatrice, Grytnes John-Arvid, Heilmann-Clausen Jacob, Boddy Lynne, Bässler Claus, Gange Alan C., Heegaard Einar, Høiland Klaus, Kirk Paul M., Krisai-Greilhüder Irmgard, Kuyper Thomas W., Kauserud Håvard (2019). Open-source data reveal how collections-based fungal diversity is sensitive to global change. *Applications in Plant Sciences* 2019 Mar; 7(3). Saatavilla (viitattu 15.5.2019): https://www.researchgate.net/publication/331691012_Open-source_data_reveal_how_collections-based_fungal_diversity_is_sensitive_to_global_change

Atlas for the End of the World (2019). Verkkosivusto. Saatavilla (viitattu 18.7.2019): http://atlas-for-the-end-of-the-world.com/world_maps_main.html

Bamber Jonathan L., Oppenheimer Michael, Kopp Robert E., Aspinall Willy P., Cooke Roger M. (2019). Ice sheet contributions to future sea-level rise from structured expert judgment. *Proceedings of the National Academy of Sciences of the United States of America*. Saatavilla (viitattu 21.05.2019): <https://www.pnas.org/content/early/2019/05/14/1817205116>

Bastin Jean-Francois, Finegold Yelana, Garcia Claude, Mollicone Danilo, Rezende Marcelo, Routh Devin, Zohner Constantin M., Crowther Thomas W. (2019). The global tree restoration potential. *Science* 05 Jul 2019. Vol. 365, Issue 6448, pp. 76-79. Saatavilla (viitattu 5.7.2019): <https://science.sciencemag.org/content/365/6448/76>

Bliss Angela C., Michael Steele Michael, Peng Ge, Meier Walter N, Dickinson Suzanne (2019). Regional variability of Arctic sea ice seasonal change climate indicators from a

passive microwave climate data record. Environmental Research Letters 14 (2019). Saatavissa (viitattu 13.5.2019): https://www.researchgate.net/publication/330097108_Regional_variability_of_Arctic_sea_ice_seasonal_change_climate_indicators_from_a_passive_microwave_climate_data_record

Boas Ingrid, Dahm Ruben, Wrathall David (2019): Grounding Big Data On Climate-Induced Human Mobility. Geographical Review 1–15. Special Issue: Fieldwork in Geography. Saatavilla (viitattu 16.5.2019): <https://onlinelibrary.wiley.com/doi/full/10.1111/gere.12355>

Braun Henrik (2015). Evaluation of Big Data Maturity Models –A Benchmarking Study to Support Big Data Maturity Assessment in Organizations. Tampereen Teknillinen Yliopisto. Diplomityö. Saatavissa (viitattu 16.7.2019): <https://trepo.tuni.fi/bitstream/handle/123456789/23016/braun.pdf>

Brönnimann Stefan, Martius Olivia, Rohr Christian, Bresch David N., Lin Kuan-Hui Elaine (2019): Historical weather data for climate risk assessment. Annals of The New York Academy of Sciences. Special Issue: Climate Sciences. 2019 Jan;1436(1):121-137. Saatavilla (viitattu 16.5.2019): <https://www.ncbi.nlm.nih.gov/pubmed/30291628>

Burillo Daniel, Chester Mikhail V., Pincetl Stephanie, Fournier Eric D., Reynac Janet (2019). Forecasting peak electricity demand for Los Angeles considering higher air temperatures due to climate change. Applied Energy. Volume 236, 15 February 2019, Pages 1-9. Saatavilla (viitattu 11.6.2019): <https://www.sciencedirect.com/science/article/pii/S0306261918317458>

Chen Chao, Kalra Ajay, Ahmad Sajjad (2019). Hydrologic responses to climate change using downscaled GCM data on a watershed scale. Journal of Water and Climate Change 10.1 63-77. 2019. Saatavilla (viitattu 22.5.2019): <https://iwaponline.com/jwcc/article/10/1/63/39055/Hydrologic-responses-to-climate-change-using>

Climate-Data.gov (2019). Saatavilla (viitattu 19.6.2019): <https://www.data.gov/climate/>

Cook John, Oreskes Naomi, Doran Peter T, Anderegg William RL, Verheggen Bart, Maibach Ed W, Carlton J Stuart, Lewandowsky Stephen, Skuce Andrew G, Green Sarah A, Nuccitelli Dana, Jacobs Peter, Richardson Mark, Winkler Bärbel, Painting Rob, Rice Ken (2016). Consensus on consensus: a synthesis of consensus estimates on human-caused global warming. Environmental Research Letters 11 (2016) 048002. Saatavilla (viitattu 5.9.2019): <https://iopscience.iop.org/article/10.1088/1748-9326/11/4/048002>

Creutzig Felix, Lohrey Steffen, Bai Xuemei, Baklanov Alexander, Dawson Richard, Dhakal Shobhakar, Lamb William F., McPhearson Timon, Minx Jan, Munoz Esteban, Walsh Brenna (2019). Upscaling urban data science for global climate solutions. *Global Sustainability*, Volume 2 2019. Saatavissa (viitattu 7.5.2019): <https://www.cambridge.org/core/journals/global-sustainability/article/upscaling-urban-data-science-for-global-climate-solutions/D2D622B43CD50A9B2FD5DF855BCC0F18>

Crowtherlabs (2019). Trees against climate change: the global restoration and carbon storage potential. Verkkosivusto. Saatavilla (viitattu 5.7.2019): <https://crowtherlab.pageflow.io/trees-against-climate-change-the-global-restoration-and-carbon-storage-potential>

Cubasch Ulrich, Kadow Christopher (2011). Global Climate Change and Aspects of Regional Climate Change in the Berlin-Brandenburg Region. *Die Erde. Journal Of The Geographical Society Of Berlin*. Vol 142, No 1-2 (2011). Saatavilla (viitattu 20.6.2019): https://www.researchgate.net/publication/265242134_Global_Climate_Change_and_Aspects_of_Regional_Climate_Change_in_the_Berlin-Brandenburg_Region

Data-Pop Alliance (2017). Big Data for Climate Resilience. Inputs for World Development Report 2016 "Digital Dividends". Saatavilla (viitattu 30.6.2019): <http://datapopalliance.org/wp-content/uploads/2017/03/WDR-Data-Pop-Alliance-input-2-Climate-Change-Resilience.pdf>

de Jong Pieter, Barreto Tarssio B., Tanajura Clemente A.S., Kouloukoui Daniel, Oliveira-Esquerre Karla P., Kiperstok Asher, Torres Ednildo Andrade (2019): Estimating the impact of climate change on wind and solar energy in Brazil using a South American regional climate model. *Renewable Energy* 141 (2019) 390-401. Saatavissa (viitattu 9.5.2019): <https://www.sciencedirect.com/science/article/pii/S096014811930391X>

Diamond Jared (2010). *Romahdus - Miten yhteiskunnat päättävät tuhoutua tai menestyä*. Terra Cognita. ISBN 9789525202847. 3 painos.

Diffenbaugh, N. S. Burke, M. (2019). Global warming has increased global economic inequality. *Proceedings of the National Academy of Sciences of the United States of America*. Saatavissa (viitattu 4.5.2019): <https://www.pnas.org/content/early/2019/04/16/1816020116>

Dietrich Helge, Wolf Thilo, Kawohl Tobias, Wehberg Jan, Kändler Gerald, Mette Tobias, Röder Arno, Böhner Jürgen (2019). Temporal and spatial high-resolution climate

data from 1961 to 2100 for the German National Forest Inventory (NFI). *Annals of Forest Science* (2019) 76: 6. Saatavilla (viitattu 13.5.2019): https://www.researchgate.net/publication/322273216_Temporal_and_spatial_high-resolution_climate_data_from_1961_to_2100_for_the_German_National_Forest_Inventory_NFI

EIA (2019). Energy Distruptions. Energy Infrastructure with Active Storms and Other Hazards. U.S. Energy Information Administration. Saatavilla (viitattu 27.6.2019): <https://www.eia.gov/special/disruptions/>

Engle Robert, Giglio Stefano, Lee Heebum, Kelly Bryan, Stroebe Johannes (2019). Hedging Climate Change News. NBER Working Paper No. 25734 Issued in April 2019. Saatavilla (viitattu 13.6.2019): https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3317570

Elinkeinoelämän Valtuuskunta (2019). Elämänmuutoksen Empijät. Julkaisu. Eva analyysi No. 73, 26.6.2019. Saatavilla (viitattu 26.6.2019): https://www.eva.fi/wp-content/uploads/2019/06/eva_analyysi_no_73-5.pdf

Faghmous James H, Kumar Vipin (2014). A Big Data Guide To Understanding Climate Change: The Case for Theory-Guided Data Science. *Big Data*. 2014 Sep 1; 2(3): 155–163. Saatavilla (viitattu 4.6.2019): <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4174912/pdf/big.2014.0026.pdf>

Financial Times (2019). The New Agenda. Capitalism. Time for a reset. Uutinen. Saatavilla (viitattu 19.9.2019): <https://aboutus.ft.com/en-gb/new-agenda/>

Fink Arlene (2019). *Conducting Research Literature Reviews*. Sage, 2010. 304 sivua. ISBN-13: 978-1544318479

Fortune (2014). Big data's biggest challenge: climate change. *Fortune* 23.4.2016. Saatavilla (viitattu 19.6.2019): <http://fortune.com/2014/06/23/big-data-climate-change-map-sea-levels/>

Fred Jonna (2018). Data Monetization – Miten organisaatio voi tuottaa liikevaihtoa datan avulla? Tampereen Teknillinen Yliopisto. Diplomityö. Saatavilla (viitattu 16.7.2019): <http://urn.fi/URN:NBN:fi:tyy-201703281232>

Global Commission On Adaptation (2019). *Adapt Now: A Global Call For Leadership On Climate Resilience*. Saatavilla (viitattu 11.9.2019): <https://cdn.gca.org/assets/2019->

[09/GlobalCommission_Report_FINAL.pdf](#)

Google Earth Engine (2019). Saatavilla (viitattu 19.6.2019): <https://earthengine.google.com/>

Google Trends (2019). Saatavilla (viitattu 14.5.2019): <https://trends.google.fi/trends/explore?date=today%205-y&q=climate%20change,big%20data,artificial%20intelligence,open%20data>

Greer John Michael (2008). The Long Descent: A User's Guide to the End of the Industrial Age. New Society Publishers; 1st Edition edition (September 1, 2008). ISBN-10: 0865716099, ISBN-13: 978-0865716094.

Guo Hua-Dong, Zhang Li, Zhu Lan-Wei (2015). Earth observation big data for climate change research. Advances in Climate Change Research, Volume 6, Issue 2, June 2015, Pages 108-117. Saatavilla (viitattu 4.6.2019): <https://www.sciencedirect.com/science/article/pii/S1674927815000519>

Hassani Hossein, Huang Xu, Silva Emmanuel (2019). Big Data and Climate Change. Big Data and Cognitive Computing. 2019, 3, 12. Saatavilla (viitattu 28.5.2019): <https://www.mdpi.com/2504-2289/3/1/12/pdf>

Haunschild Robin, Bornmann Lutz, Marx Werner (2016). Climate Change Research in View of Bibliometrics. PLoS ONE 11(7): e0160393. doi:10.1371/journal.pone.0160393. Saatavilla (viitattu 5.9.2019): <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0160393>

Heikura Jarkko (2018). Financial sentiment analysis from big data. Tampereen Yliopisto, Johtamiskorkeakoulu. Diplomityö. Saatavilla (metadata, viitattu 16.7.2019): <https://trepo.tuni.fi/handle/10024/103382>

Heinonen Kristiina (2016). Avoimen datan hyödyntämisen haasteet globaalien organisaatioiden big data analytiikassa. Tampereen teknillinen yliopisto, diplomityö. Saatavilla (viitattu 7.6.2019): <https://dspace.cc.tut.fi/dpub/bitstream/handle/123456789/23843/Heinonen.pdf>

Helsingin Sanomat (2019a). Talouslehti Financial Times vetoaa: Kapitalistinen järjestelmä on nollattava. Uutinen 18.9.2019. Saatavilla (viitattu 19.9.2019): <https://www.hs.fi/talous/art-2000006243152.html>

Helsingin Sanomat (2019b). Suuryritysten johtajat julistavat poikkeuksellisessa kannanotossa muunkin kuin voiton olevan tärkeää. Uutinen 20.8.2019. Saatavilla (viitattu 19.9.2019): <https://www.hs.fi/talous/art-2000006210304.html>

Herala Antti (2018). Benefits from Open Data: Barriers to Supply and Demand of Open Data in Private Organizations. Lappeenrantaan Teknillinen Yliopisto, Väitöskirja. Saatavissa (viitattu 18.7.2019): <http://lutpub.lut.fi/handle/10024/158560>

Hildén Mikael, Groundstroem Fanny, Carter Timothy R., Halonen Mikko, Perrels Adrian, Gregow Hilppa (2016). Ilmastomuutoksen heijastevaikutukset Suomeen. Valtioneuvoston selvitys- ja tutkimustoiminnan julkaisusarja 46/2016. Saatavissa (viitattu 5.5.2019):

https://tietokayttoon.fi/documents/10616/2009122/46_Ilmastomuutoksen+heijastevaikutukset+Suomeen/

Hurlimann Anna Catherine, Warren-Myers Georgia, Browne Geoffrey R. (2019). Is the Australian construction industry prepared for climate change? Building and Environment 153 (2019) 128–137. Saatavilla (viitattu 11.6.2019): <https://www.sciencedirect.com/science/article/pii/S0360132319301039>

IBM (2019). The Four V's of Big Data. Infographics & Animations. Saatavilla (viitattu 4.9.2019): <https://www.ibmbigdatahub.com/infographic/four-vs-big-data>

Ilmatieteen Laitos (2019). Tulvakeskuksen sivut. Saatavilla (viitattu 22.5.2019): <https://ilmatieteenlaitos.fi/tulvakeskus>

IPCC (2007). Historical Overview of Climate Change. AR4 Climate Change 2007. Saatavilla (viitattu 5.9.2019): <https://www.ipcc.ch/report/ar4/wg1/historical-overview-of-climate-change-science/>

IPCC (2019). About - IPCC. Saatavilla (viitattu 30.6.2019): <https://www.ipcc.ch/about/>

Jang S. Mo, Hart P. Sol (2015). Polarized frames on “climate change” and “global warming” across countries and states: Evidence from Twitter big data. Global Environmental Change. Volume 32, May 2015, Pages 11-17. Saatavilla (viitattu 25.6.2019): <https://www.sciencedirect.com/science/article/abs/pii/S0959378015000291>

Kentala-Lehtonen Johanna (2019). Climate Change as Problem of Direction and Pace of Transition : Large Finnish Business Actors' Identity, Interests, and Political Response

Strategies to Climate Politics. Tampereen Yliopisto, väitöskirja. Kansainvälinen politiikka, johtamisen ja talouden tiedekunta. Saatavilla (viitattu 18.7.2019): <https://trepo.tuni.fi/handle/10024/104759>

Kinnunen Tiia (2016). Vihreä tavarankuljetuslogistiikka Suomessa. Haaga-Helia Ammattikorkeakoulu Oy. Liiketalouden koulutusohjelma. Opinnäytetyö. Saatavilla (viitattu 26.6.2019): https://www.theseus.fi/bitstream/handle/10024/106675/Opinnaytetyo_TiiaKinnunen.pdf

Knüsel Benedikt, Zumwald Marius, Baumberger Christoph, Hadorn Gertrude Hirsch, Fischer Erich M., Bresch David N., Knutti Reto (2019). Applying big data beyond small problems in climate research. Nature Climate Change volume 9, pages 196–202 (2019). Vain tiivistelmä saatavilla (viitattu 18.7.2019): <https://www.nature.com/articles/s41558-019-0404-1>

Kronholm Kristiina Eva Amanda (2018). Kaupunkisuunnittelun digitalisaatio. Jyväskylän yliopisto, Informaatioteknologinen tiedekunta, 2018. Tietojärjestelmätiede, Kandidaatintutkielma. Saatavilla (viitattu 26.6.2019): <https://jyx.jyu.fi/bitstream/handle/123456789/57910/URN%3ANBN%3Afi%3Aju-201805092512.pdf>

König Christian, Weigelt Patrick, Schrader Julian, Taylor Amanda, Kattge Jens, Kreft Holger (2019). Biodiversity data integration—the significance of data resolution and domain. PLoS Biol 17(3): e3000183. Saatavilla (viitattu 20.5.2019): <https://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.3000183>

Lassinantti Josefin, Ståhlbröst Anna, Runardotter Mari (2019). Relevant social groups for open data use and engagement. Government Information Quarterly 36 (2019) 98–111. Saatavilla (viitattu 28.5.2019): <https://www.sciencedirect.com/science/article/pii/S0740624X18300959>

Lavonen Emilia (2013). Maaseutu sanomalehtien ilmastonmuutoskirjoituksissa. Helsingin yliopisto, Matemaattis-luonnontieteellinen tiedekunta, Geotieteiden ja maantieteen laitos. Pro gradu -tutkielma. Saatavissa (viitattu 4.6.2019): <https://helda.helsinki.fi/handle/10138/230476>

Lehikoinen Emmi Sofia (2014). Sosiaalinen mediatiedeorganisaatioiden ilmastonmuutosviestinnän foorumina. Helsingin Yliopisto, Valtiotieteellinen tiedekunta, Sosiaalitieteiden laitos. Pro gradu. Saatavilla (viitattu 16.7.2019): https://helda.helsinki.fi/bitstream/handle/10138/45184/ProGradu_LehikoinenEmmi.pdf

Li Francis G.N., Bataille Chris, Pye Steve, O'Sullivan Aidan (2019): Prospects for energy economy modelling with big data: Hype, eliminating blind spots, or revolutionising the state of the art?. *Applied Energy*, Volume 239, 1 April 2019, Pages 991-1002. Saatavilla (viitattu 15.5.2019): <https://www.sciencedirect.com/science/article/pii/S0306261919302922>

Luonto-Liitto (2019). Luonto-Liiton kevätseuranta. Saatavissa (viitattu 7.5.2019): <http://www.kevatseuranta.fi/>

Maliniemi Tuija (2018). Decadal Time-Scale Vegetation Changes At High Latitudes. Responses To Climatic And Non-Climatic Drivers. Väitöskirja, Oulun Yliopisto. Saatavilla (viitattu 5.6.2019): <http://jultika.oulu.fi/files/isbn9789526220123.pdf>

Mentula Minna (2018). Ilmasto, energiaturvallisuus ja talouskasvu eduskunnan diskursseissa. Analyysissä energia- ja ilmastopoliittinen lähetekeskustelu syksyllä 2016. Johtamiskorkeakoulu, Ympäristöpolitiikka ja aluetiede, Pro Gradu -tutkielma ja seminaari 2018. Saatavilla (viitattu 4.6.2019): <https://tampub.uta.fi/bitstream/handle/10024/104895/1546515672.pdf>

Michetti Melania, Pinar Mehmet (2019). Forest Fires Across Italian Regions and Implications for Climate Change: A Panel Data Analysis. *Environmental and Resource Economics* (2019) 72:207–246. Saatavilla (viitattu 20.5.2019): <https://link.springer.com/article/10.1007/s10640-018-0279-z>

Muszynski Grzegorz, Kashinath Karthik, Kurlin Vitaliy, Wehner Michael, Prabhat (2019). Topological data analysis and machine learning for recognizing atmospheric river patterns in large climate datasets. *Geoscientific Model Development* 12, 613–628, 2019. Saatavilla (viitattu 28.5.2019): <https://www.geosci-model-dev.net/12/613/2019/>

Mäkinen Katja (2019). Climate-induced variability in northern Baltic Sea zooplankton - Assessing driving forces and effects on higher trophic levels. Väitöskirja, Turun Yliopisto. Saatavissa (viitattu 5.6.2019): <https://www.utupub.fi/handle/10024/147418>

Neukom Raphael, Steiger Nathan, Gómez-Navarro Juan José, Wang Jianghao, Werner Johannes P. (2019). No evidence for globally coherent warm and cold periods over the preindustrial Common Era. *Nature. Letter* 24 Julu 2019. *Nature* volume 571, pages 550–554 (2019). Saatavilla (viitattu 25.7.2019): <https://www.nature.com/articles/s41586-019-1401-2>

Okoli C (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. Working Papers on Information Systems. ISSN 1535-6078. Saatavilla (viitattu 3.10.2019): <https://pdfs.semanticscholar.org/31dc/753345d5230e421ea817dd7dcdd352e87ea2.pdf>

PAGES (2019). Past Global Changes. Nettisivusto. Saatavilla (viitattu 25.7.2019): <http://www.pastglobalchanges.org/>

Penteriani Vincenzo, Zarzo-Arias Alejandra, Novo-Fernández Alís, Bombieri Giulia, López-Sánchez Carlos A. (2019). Responses of an endangered brown bear population to climate change based on predictable food resource and shelter alterations. *Global Change Biology*, Volume 25, Issue 3. Saatavilla (viitattu 12.6.2019): <https://onlinelibrary.wiley.com/doi/abs/10.1111/gcb.14564>

Reichstein Markus, Camps-Valls Gustau, Stevens Bjorn, Jung Martin, Denzler Joachim, Carvalhais Nuno, Prabhat (2019). Deep learning and process understanding for data-driven Earth system science. *Nature*, vol 566, February 2019. Saatavissa (viitattu 8.5.2019): <https://www.nature.com/articles/s41586-019-0912-1>

Reside April E., Critchell Kay, Crayn Darren M., Goosem Miriam, Goosem Stephen, Hoskin Conrad J., Sydes Travis, Vanderduys Eric P., Pressey Robert L. (2019). Beyond the model: expert knowledge improves predictions of species' fates under climate change. *Ecological Applications*, Volume 29, Issue 1, January 2019. Saatavilla (viitattu 13.6.2019): <https://esajournals.onlinelibrary.wiley.com/doi/abs/10.1002/eap.1824>

Salmivaara Aura (2015). Spatial Vulnerability Assessments for Water Resources Management - Cases from major Asian river basins with a focus on spatial unit of analysis and the use of big and open data - Paikkatietoa hyödyntävät haavoittuvuusarvioinnit vesivarojen hallinnassa: Tapaustutkimuksia Aasian merkittävistä jokivesistöistä keskittyen analyysiyksikköön sekä laajan ja avoimen datan käyttöön. Aalto-yliopisto. Väitöskirja. Saatavissa (viitattu 5.5.2019): <https://aaltodoc.aalto.fi/handle/123456789/17335>

Sathiaraj D., X. Huang X., Chen J. (2019). Predicting climate types for the Continental United States using unsupervised clustering techniques. *Environmetrics*. 2019;30:e2524. Saatavilla (viitattu 20.5.2019): <https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2524>

Schnase John L., Lee Tsengdar J., Mattmann Chris A., Lynnes Christopher S., Cinquini Luca, Ramirez Paul M., Hart Andre F., Williams Dean N., Waliser Duane, Rinsland Pamela, Webster W. Philip, Duffy Daniel Q., McNerney Mark A., Tamkin Glenn S., Potter

Gerald L., Carrier Laura (2015). Big Data Challenges in Climate Science. NASA, GRSM-2015-00014.R2. Saatavilla (viitattu 19.6.2019): <https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20160002946.pdf>

Schneider Birgit, Walsh Lynda (2019). The politics of zoom: Problems with downscaling climate visualizations. Geography and Environment. Volume 6, Issue 1. January-June 2019. Saatavilla (viitattu 3.7.2019): https://www.researchgate.net/publication/330742626_The_politics_of_zoom_Problems_with_downscaling_climate_visualizations

Science (2019). Earth scientists plan to meld massive databases into a ‘geological Google’. Science Feb. 26, 2019. Saatavilla (viitattu 22.6.2019): <https://www.sciencemag.org/news/2019/02/earth-scientists-plan-meld-massive-databases-geological-google>

Shukla Roopam, Agarwal Ankit, Sachdeva Kamna, Kurths Juergen, Joshi P.K. (2019). Climate change perception: an analysis of climate change and risk perceptions among farmer types of Indian Western Himalayas. Climatic Change (2019) 152:103–119. Saatavilla (viitattu 14.6.2019): <https://link.springer.com/article/10.1007/s10584-018-2314-z>

Siemens (2019): Virtuaalivoimalaitos – Kiinteistöt hyötymään sähkömarkkinoista. Saatavilla (viitattu 15.5.2019): <https://new.siemens.com/fi/fi/yhtio/ajankohtaiset-teemat/alykas-infrastruktuuri/kiinteistoejen-virtuaalivoimalaitoshanke-lisaa-omistajien-tuot.html>

Singh Sanjay Kumar, El-Kassar Abdul-Nasser (2019). Role of big data analytics in developing sustainable capabilities. Journal of Cleaner Production 213 (2019) 1264–1273. Saatavilla (viitattu 14.5.2019): <https://www.sciencedirect.com/science/article/pii/S0959652618339155>

SmartDataCollective (2019). Can Big Data Help Save The Environment? Ryan Kh, February 8, 2019. Saatavilla (viitattu 28.6.2019): <https://www.smartdatacollective.com/can-big-data-help-save-environment/>

Srinivasana Rengarajan, Giannikasb Vaggelis, Kumara Mukesh, Guyota Renaud, McFarlanea Duncan (2019). Modelling food sourcing decisions under climate change: A data-driven approach. Computers & Industrial Engineering 128 (2019) 911–919. Saatavissa (viitattu 8.5.2019): <https://www.sciencedirect.com/science/article/abs/pii/S036083521830528X>

Steiger Robert, Scott Daniel, Abegg Bruno, Pons Marc, Aall Carlo (2019). A critical review of climate change risk for ski tourism. *Current Issues in Tourism*, 2019. Vol. 22, No. 11, 1343–1379. Saatavilla (viitattu 11.6.2019): <https://www.tandfonline.com/doi/full/10.1080/13683500.2017.1410110>

Stephens-Davidowitz Seth (2017). *Everybody Lies: Big Data, New Data, and What the Internet Can Tell Us About Who We Really Are*. Dey Street Books (May 9, 2017). ISBN-10: 0062390856, ISBN-13: 978-0062390851.

Sudmanns Martin, Tiede Dirk, Lang Stefan, Bergstedt, Trost Georg, Augustin Hannah, Baraldi Andrea, Blaschke Thomas (2019): Big Earth data: disruptive changes in Earth observation data management and analysis?. *International Journal of Digital Earth*. Saatavissa (viitattu 7.5.2019): <https://www.tandfonline.com/doi/full/10.1080/17538947.2019.1585976>

Suojelupoliisi (2018). Juhlavuosisikirja 2018. Suojelupoliisin julkaisuja. Saatavilla (viitattu 28.5.2019): <https://www.supo.fi/julkaisut/esitteet>

Syrjärinne Paula (2016). *Urban Traffic Analysis with Bus Location Data*. Tampereen Yliopisto. Väitöskirja. Saatavissa (viitattu 5.5.2019): <http://tampub.uta.fi/handle/10024/98613>

Tabvizexplorer (2018). *Sentiment Analysis using R and Twitter*. Mithun Desai. Saatavilla (viitattu 13.6.2019): <https://tabvizexplorer.com/sentiment-analysis-using-r-and-twitter/>

Tainter Joseph A. (1998). *The Collapse of Complex Societies*. Cambridge University Press. SBN 0 521 340926.

Tekes (2017). *Tulevaisuuden energia 2030...2050 Taustaraportti*. Pia Salokoski. Tekesin katsaus 332/2017. Helsinki 2017. Saatavilla (viitattu 28.6.2019): https://tem.fi/documents/1410877/2772829/332_2017_Tulevaisuuden+energia_2030_2050.pdf/4f1c0ec0-58fc-4c1c-9297-7f90ac01615b

Tekniikan Maailma (2019a). Venäjällä olisi huima mahdollisuus 151 miljoonan hehtaarin metsittämiseen – Tutkijat esittävät valtavaa maapallon metsittämistä ratkaisuksi ilmastonmuutoksen hillitsemiseen. Uutinen 5.7.2019. Saatavilla (viitattu 5.7.2019): <https://tekniikanmaailma.fi/venajalla-olisi-huima-mahdollisuus-151-miljoonan->

[hehtaarin-metsittämiseen-tutkijat-esittavat-valtavaa-maapallon-metsittamista-ratkaisuksi-ilmastonmuutoksen-hillitsemiseen/](#)

Tekniikan Maailma (2019b). 1900-luvulla tapahtunut ilmastonmuutos on voimakkainta 2000 vuoteen, kertoo tutkimus – Esimerkiksi keskiajan ”pieni jääkausi” oli siihen verrattuna täysin erilainen ilmiö. Uutinen 25.7.2019. Saatavilla (viitattu 25.7.2019): <https://tekniikanmaailma.fi/1900-luvulla-tapahtunut-ilmastonmuutos-on-voimakkainta-2000-vuoteen-kertoo-tutkimus-esimerkiksi-keskiajan-pieni-jaakausi-oli-siihen-verrattuna-taysin-erilainen-ilmio/>

Tekniikka ja Talous (2018): Pörssin sekuntikaupassa osake pidetään vain häviävän hetken – algoritmi nykymeklarille kuin "saha metsurille". Uutinen 16.4.2018. Saatavilla (viitattu 30.6.2019): <https://www.tekniikkatalous.fi/uutiset/porssin-sekuntikaupassa-osake-pidetaan-vain-haviavan-hetken-algoritmi-nykymeklarille-kuin-saha-metsurille/ba804c01-b6ad-3b27-a538-7fc4f40cd70a>

Tekniikka ja Talous (2019). Uutta tutkimustietoa paljon spekuloidusta aiheesta: ilmastonmuutos todella aiheuttaa konflikteja ja pakottaa ihmisiä muuttamaan. Uutinen 23.1.2019 13:33. Saatavilla (viitattu 16.7.2019): <https://www.tekniikkatalous.fi/uutiset/uutta-tutkimustietoa-paljon-spekuloidusta-aiheesta-ilmastonmuutos-todella-aiheuttaa-konflikteja-ja-pakottaa-ihmisia-muuttamaan/b18d8573-9453-3577-aa11-4b5b3e28bcf6>

The Guardian (2019). 'No doubt left' about scientific consensus on global warming, say experts. Uutinen 24.7.2019. Saatavilla (viitattu 25.7.2019): <https://www.theguardian.com/science/2019/jul/24/scientific-consensus-on-humans-causing-global-warming-passes-99>

Tiampo Kristy F. , Kazemian Javad, Ghofrani Hadi, Kropivnitskaya Yelena, Michel Gero (2019). Insights into seismic hazard from big data analysis of ground motion simulations. International Journal of Safety and Security Engineering. Vol. 9, No. 1 (2019) 1–12. Saatavilla (viitattu 22.5.2019): <https://www.witpress.com/elibrary/sse-volumes/9/1/2367>

Tiede-lehti (2010). Ilmastoskeptikot taitavat markkinoinnin. 27.4.2010. Tiede lehti 3/2010. Saatavilla (viitattu 26.6.2019): https://www.tiede.fi/artikkeli/jutut/artikkelit/ilmastoskeptikot_taitavat_markkinoinnin

Tivi (2019). Julkinen sektori istuu datan päällä – ”koordinointi tulisi olla yhdellä vastuuministerillä”. Uutinen 15.5.2019. Saatavilla (viitattu 16.5.2019):

<https://www.tivi.fi/uutiset/julkinen-sektori-istuu-datan-paalla-koordinointi-tulisi-olla-yhdella-vastuuministerilla/cc920fff-1b04-417f-a9a0-e44812a88dfb>

TobaccoWatcher (2019). Saatavilla (viitattu 27.6.2019):
<https://tobaccowatcher.globaltobaccocontrol.org/analyses/>

Toivonen Tuomas (2016). Maanteiden kuivatusjärjestelmien ja kuivatukseen liittyvän tiedonhallinnan kehitystarpeet ilmaston muuttuessa. Tampereen Teknillinen Yliopisto. Diplomityö. Saatavissa (viitattu 5.5.2019):
<https://dspace.cc.tut.fi/dpub/bitstream/handle/123456789/23624/Toivonen.pdf>

Uudenkaupungin Sanomat (2019). Hätäkeskuksen uusi tietojärjestelmä keskustelee paremmin yli maakuntarajojen. Uutinen 3.4.2019. Saatavilla (viitattu 16.5.2019):
<https://www.uudenkaupunginsanomat.fi/2019/04/hatakeskuksen-uusi-tietojarjarjestelma-keskustelee-paremmi-yli-maakuntarajojen/>

World Economic Forum (2019). The Global Risks Report 2019. 14th Edition. Saatavissa (viitattu 3.5.2019):
http://www3.weforum.org/docs/WEF_Global_Risks_Report_2019.pdf

Yang Chaowei, Manzhu Yu Manzhu, Li Yun, Hu Fei, Jiang Yongyao, Liu Qian, Sha Dexuan, Xu Mengchao, Gu Juan (2019). Big Earth data analytics: a survey. Big Earth Data 2019, Vol. 3, No. 2, 83-107. Saatavilla (viitattu 2.7.2019):
<https://doi.org/10.1080/20964471.2019.1611175>

Yle (2008). Tutkimus: Ilmastomuutos lisää astmaa ja allergioita. Uutinen 1.11.2008. Saatavilla (viitattu 21.5.2019): <https://yle.fi/uutiset/3-5850440>

Yle (2009). Ilmastomuutos ajanut satoja tuhansia afrikkalaisia pakolaisiksi. Uutinen 23.10.2009. Saatavilla (viitattu 16.7.2019): <https://yle.fi/uutiset/3-5904715>

Yle (2011). IPCC varoittaa sään vaarallisista ääri-ilmiöistä. Uutinen 2.11.2011. Saatavilla (viitattu 22.5.2019): <https://yle.fi/uutiset/3-5446182>

Yle (2015a). Myytti tiiviistä ekokaupungista – eli kertomus siitä miten ajatusharhasta tuli vallitseva totuus. Pasi Toiviainen. 11.9.2015. Saatavilla (viitattu 26.6.2019): <https://yle.fi/aihe/artikkeli/2015/08/31/myytti-tiiviista-ekokaupungista-eli-kertomus-siita-miten-ajatusharhasta-tuli>

Yle (2015b). Järven pohjassa lojuu ikuinen arkisto. Uutinen 3.12.2015. Saatavilla (vii-

tattu 30.6.2019): <https://yle.fi/aihe/artikkeli/2015/12/02/jarven-pohjassa-lojuu-ikuinen-arkisto>

Yle (2017a). Etelä-Eurooppaa piinaa ennätysellinen helleaalto – Turistikohteita suljettu. Uutinen 14.7.2017. Saatavilla (viitattu 14.5.2019): <https://yle.fi/uutiset/3-9722904>

Yle (2017b). Tekoäly alkaa takoa rahaa sijoittajille – kone lukee kaikki maailman vuosikertomukset muutamassa päivässä. Uutinen 16.11.2017. Saatavilla (viitattu 13.6.2019): <https://yle.fi/uutiset/3-9933010>

Yle (2018a). Entä jos jokainen tekosi tallentuisi kameralle ja sinut pisteytettäisiin kansalaisena? Kiinassa se on pian totta. Uutinen 28.3.2018. Saatavilla (viitattu 21.5.2019): <https://yle.fi/uutiset/3-10135093>

Yle (2018b). 3D-teknologia mullistaa tulvasuojelua – Suomessa kehitetty pistepilviteknologia tekee tulvien seurannasta nykyistä tehokkaampaa. Uutinen 9.5.2018. Saatavilla (viitattu 16.5.2019): <https://yle.fi/uutiset/3-10196814>

Yle (2018c). Katso animaatiosta, miten Arktis sulaa – Mitä tapahtuu hauralle eliöstölle ja pohjoisen ilmastolle? Uutinen 20.3.2018. Saatavilla (viitattu 14.5.2019): <https://yle.fi/uutiset/3-10117709>

Yle (2018d). "Tällaisessa paikassa maanjäristys ei normaalisti synnytä tsunamia" – Palun kaupunki jäi hyökyaallon alle, kun pitkä lahti nostatti hyökyaallon jopa kuusimetriseksi. Uutinen 1.10.2018. Saatavilla (viitattu 22.5.2019) <https://yle.fi/uutiset/3-10432657>

Yle (2018e). Espoossa halutaan arvioida kaikkien kaupungin päätösten ilmastovaikutukset. Uutinen 16.10.2018. Saatavilla (viitattu 30.6.2019): <https://yle.fi/uutiset/3-10458605>

Yle (2018f). Pori ja Rovaniemi vesistöjen merkittävimmät tulvariskialueet. Uutinen 20.12.2018. Saatavilla (viitattu 22.5.2019): <https://yle.fi/uutiset/3-10566241>

Yle (2018g). Suomalaistutkijan animaatio ilmastomuutoksesta kohahduttaa jälleen maailmalla – "Ei näytä, että ilmaston lämpeneminen olisi vähenemässä". Uutinen 30.8.2018. Saatavilla (viitattu 20.5.2019): <https://yle.fi/uutiset/3-10377870>

Yle (2018h). Maailman ilmatieteen järjestön johtaja: "Suomi ei ole ilmastomuutoksen

voittaja" – kasvukausi voi pidentyä, mutta haitat ylittävät hyödyt. Uutinen 14.8.2018. Saatavilla (viitattu 8.5.2019): <https://yle.fi/uutiset/3-10349944>

Yle (2019a). Muovi on nyt arkkipaha, seuraavat megatrendit ehkä ruoka ja liikkuminen – miten kuluttaja erottaa viherpesun aidosta vastuullisuudesta? Uutinen 7.3.2019. Saatavilla (viitattu 14.5.2019): <https://yle.fi/uutiset/3-10674467>

Yle (2019b). Tutkimus: Siirtyminen uusiutuvaan energiaan on mahdollista, turvallista ja edullisempaa kuin fossiiliset ja ydinvoima. Uutinen 12.4.2019. Saatavilla (viitattu 14.5.2019): <https://yle.fi/uutiset/3-10734977>

Yle (2019c). Huoli ilmastonmuutoksesta on vallannut jo pienemmätkin suomalaisyritykset – "Työmailla vaaditaan, että meidän ympäristöasiat on kunnossa". Uutinen 22.5.2019. Saatavilla (viitattu 28.5.2019): <https://yle.fi/uutiset/3-10795759>

Yle (2019d). Uusiutuva polttoaine on Nesteen kultakaivos – Etumatka alkaa olla niin pitkä, että maailmasta ei löydy kunnan vastusta suomalaisyhtiölle. Uutinen 6.2.2019. Saatavilla (viitattu 21.5.2019): <https://yle.fi/uutiset/3-10633202>

Yle (2019e). Surua, epätoivoa, pelkoa – ilmastoahdistus rajoittaa nuorten tulevaisuuden suunnitelmia. Uutinen 3.5.2019. Saatavilla (viitattu 21.5.2019): <https://yle.fi/uutiset/3-10754435>

Yle (2019f). "Turhaa viherpiipertäjien hömpötystä" – kaikki eivät halua tietää ostosten sa hiilijalanjälkeä, ja tutkija kertoo miksi. Uutinen 22.5.2019. Saatavilla (viitattu 22.5.2019): <https://yle.fi/uutiset/3-10788214>

Yle (2019g). Ilmastonmuutos vaarantaa maaeläimiä mutta merissä sukupuuton uhka on kaksinkertainen. Uutinen 26.4.2019. Saatavilla (viitattu 20.5.2019): <https://yle.fi/uutiset/3-10756321>

Yle (2019h). Ilmastonmuutoksen riskeihin on Suomessa valmistauduttu puutteellisesti – osa yrityksistä tekee silti jo hyvää bisnestä. Uutinen 8.4.2019. Saatavilla (viitattu 16.5.2019): <https://yle.fi/uutiset/3-10722209>

Yle (2019i). Ruotsalaiset sytyttävät maastopaloja kertakäyttögrilleillään, joten kauppa- ketjut eivät pian myy niitä – Suomessa kaupat luottavat grillaajiin. Uutinen 24.4.2019. Saatavilla (viitattu 20.5.2019): <https://yle.fi/uutiset/3-10752782>

Yle (2019j). USA:n Keskilännen maataloutta uhkaa katovuosi – Sateiden määrä kasvaa

jatkuvasti ilmastonmuutoksen takia. Uutinen 7.6.2019. Saatavilla (viitattu 14.6.2019): <https://yle.fi/uutiset/3-10821347>

Yle (2019k). Tutkijat varoittavat: Merenpinta voi nousta jopa yli kaksi metriä. Uutinen 21.5.2019. Saatavilla (viitattu 21.5.2019): <https://yle.fi/uutiset/3-10793193>

Yle (2019l). Mosambikin tulvavuhrien määrä noussut 38:aan –Tuhoalueella tulvia ja rankkasateita. Uutinen 29.4.2019 . Saatavilla (viitattu 22.5.2019): <https://yle.fi/uutiset/3-10759957>

Yle (2019m). Sitra: Punainen liha pannaan ja omat autot vapaa-ajalla kotiin – ilman ennäkemättömän kovia tekoja ilmastotavoite karkaa Suomessa. Uutinen 16.5.2019. Saatavilla (viitattu 20.5.2019): <https://yle.fi/uutiset/3-10784895>

Yle (2019n). Tutkija: "Ilmastonmuutoksella voi olla hurjia vaikutuksia arktiseen luontoon" - Lisääntyminen epäonnistui myskihäristä pölyttäjäkärpäsiin. Uutinen 14.04.2019. Saatavilla (viitattu 15.5.2019): <https://yle.fi/aihe/artikkeli/2019/04/14/tutkija-ilmastonmuutoksella-voi-olla-hurjia-vaikutuksia-arktiseen-luontoon>

Yle (2019o). FT: Sään ääri-ilmiöt lisäsivät energiankulutusta – "Siitä voi syntyä todella ikävä noidankehä". Uutinen 11.6.2019. Saatavilla (viitattu 12.6.2019): <https://yle.fi/uutiset/3-10826962>

Yle (2019p). Kasvien sukupuuttoaalto kauhistuttaa tutkijoita – nopeutunut 500-kertaiseksi. Uutinen 11.6.2019. Saatavilla (viitattu 12.6.2019): <https://yle.fi/uutiset/3-10825261>

Yle (2019q). Tämä aggressiivinen merenpohjan vieraslaji on levinnyt jo Perämerelle asti – Mustatäplätokkoa ei voi pysäyttää, mutta olisiko siitä ruokakalaksi? Uutinen 12.6.2019. Saatavilla (viitattu 12.6.2019): <https://yle.fi/uutiset/3-10825736>

Yle (2019r). Mullistava tutkimustulos: Puiden istuttaminen on tehokkaampi ja halvempi kuin mikään muu ilmastonmuutoksen torjuntakeino. Uutinen 5.7.2019. Saatavilla (viitattu 5.7.2019): <https://yle.fi/uutiset/3-10863916>

Yle (2019s). Puolet Yhdysvalloista hikoilee kosteassa lämpöaallossa – tänä kesänä poikkeuksellisen kuumaa maailmanlaajuisesti. Uutinen 21.7.2019. Saatavilla (viitattu 25.7.2019): <https://yle.fi/uutiset/3-10887486>

Yle (2019t). Ruotsissa testataan ratkaisua, joka mullistaisi yhteiskunnan ja romauttaisi

päästöt – ja idea voi tulla käyttöön ensimmäisenä Suomessa. Uutinen 8.9.2019. Saatavilla (viitattu 10.9.2019): <https://yle.fi/uutiset/3-10942131>

Yle (2019u). Raportti: Ilmastomuutokseen sopeutuminen maksaa 1 800 miljardia dollaria – Raha tulisi moninkertaisena takaisin. Uutinen 10.9.2019. Saatavilla (viitattu 11.9.2019): <https://yle.fi/uutiset/3-10964952>

Yle (2019v). Uusi muistutus ihmiskunnan kohtalonhetkistä – YK-arvio: Luonnon tuhoutuminen ja rahan keskittyminen estämässä maailman tilan parantamisen. Uutinen 11.9.2019. Saatavilla (viitattu 12.9.2019): <https://yle.fi/uutiset/3-10958075>

Yle (2019w). Maailman vuotuinen ylikulutuspäivä jälleen ennätysaikaisin – Suomen kulutus yli keskiarvon. Uutinen 26.6.2019. Saatavilla (viitattu 27.6.2019): <https://yle.fi/uutiset/3-10848628>

Wang Tianni, Qua Zhuohua, Yang Zaili, Nichola Timothy, Dimitriu Delia, Clarke Geoff, Bowden Daniel (2019). How can the UK road system be adapted to the impacts posed by climate change? By creating a climate adaptation framework. Transportation Research Part D: Transport and Environment. 7 March 2019. Saatavilla (viitattu 11.6.2019): <https://www.sciencedirect.com/science/article/pii/S1361920918309799>

Williamson Craig E, Neale Patric J., Hylander Samuel, Rose Kevin C., Figueroa Félix L., Robinson Sharon A., Häder Donat-P., Wängberg Sten-Åke, Worrest Robert C. (2019). The interactive effects of stratospheric ozone depletion, UV radiation, and climate change on aquatic ecosystems. Photochemical & Photobiological Sciences 2019, 18, 717–746. Saatavilla (viitattu 11.6.2019): <https://pubs.rsc.org/en/content/articlelanding/2019/pp/c8pp90062k#!divAbstract>